

Cover's Open Problem: “The Capacity of the Relay Channel”

Xiugang Wu, Leighton Pate Barnes and Ayfer Özgür

Abstract

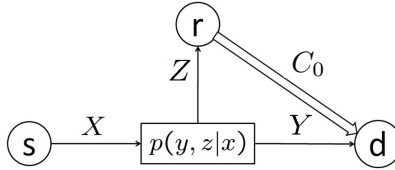
Consider a memoryless relay channel, where the channel from the relay to the destination is an isolated bit pipe of capacity C_0 . Let $C(C_0)$ denote the capacity of this channel as a function of C_0 . What is the critical value of C_0 such that $C(C_0)$ first equals $C(\infty)$? This is a long-standing open problem posed by Cover and named “The Capacity of the Relay Channel,” in *Open Problems in Communication and Computation*, Springer-Verlag, 1987. In this paper, we answer this question in the Gaussian case and show that $C(C_0)$ can not equal to $C(\infty)$ unless $C_0 = \infty$, regardless of the SNR of the Gaussian channels, while the cut-set bound would suggest that $C(\infty)$ can be achieved at finite C_0 . Our approach is geometric and relies on a strengthening of the isoperimetric inequality on the sphere by using the Riesz rearrangement inequality.

I. PROBLEM SETUP AND MAIN RESULT

In 1987, Thomas M. Cover asked the following question which he called “The Capacity of the Relay Channel” [2].

A. The Capacity of the Relay Channel¹

Consider the following seemingly simple discrete memoryless relay channel: Here Z and Y are condi-



tionally independent and conditionally identically distributed given X , that is, $p(z, y|x) = p(z|x)p(y|x)$. Also, the channel from Z to Y does not interfere with Y . A $(2^{nR}, n)$ code for this channel is a map $X^n : [1 : 2^{nR}] \rightarrow \mathcal{X}^n$, a relay function $f_n : \mathcal{Z}^n \rightarrow [1 : 2^{nC_0}]$ and a decoding function $g_n : \mathcal{Y}^n \times [1 : 2^{nC_0}] \rightarrow [1 : 2^{nR}]$. The probability of error is given by

$$P_e^{(n)} = \Pr(g_n(Y^n, f_n(Z^n)) \neq M),$$

where the message M is uniformly distributed over $[1 : 2^{nR}]$ and

$$p(m, y^n, z^n) = 2^{-nR} \prod_{i=1}^n p(y_i | x_i(m)) \prod_{i=1}^n p(z_i | x_i(m)).$$

Let $C(C_0)$ be the supremum of achievable rates R for a given C_0 , that is, the supremum of the rates R for which $P_e^{(n)}$ can be made to tend to zero. We note the following facts:

The work was supported in part by NSF award CCF-1514538 and by the Center for Science of Information (CSoI), an NSF Science and Technology Center, under grant agreement CCF-0939370. This paper was presented in part at the 2015 Allerton Conference on Communication, Control, and Computing [1].

X. Wu, L. P. Barnes and A. Özgür are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA (e-mail: x23wu@stanford.edu; lpb@stanford.edu; aozgur@stanford.edu).

¹This subsection is taken verbatim from [2] with a few notation changes.

1. $C(0) = \sup_{p(x)} I(X; Y)$.
2. $C(\infty) = \sup_{p(x)} I(X; Y, Z)$.
3. $C(C_0)$ is a nondecreasing function of C_0 .

What is the critical value of C_0 such that $C(C_0)$ first equals $C(\infty)$?

B. Main Result

In this paper, we answer this long-standing open question in the Gaussian case. In particular, consider the symmetric Gaussian relay channel as depicted in Fig. 1, where

$$\begin{cases} Z = X + W_1 \\ Y = X + W_2 \end{cases}$$

with the transmitted signal being constrained to average power P , i.e.²,

$$\frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \|X^n(m)\|^2 \leq nP, \quad (1)$$

and $W_1, W_2 \sim \mathcal{N}(0, N)$ representing Gaussian noises that are independent of each other and X .

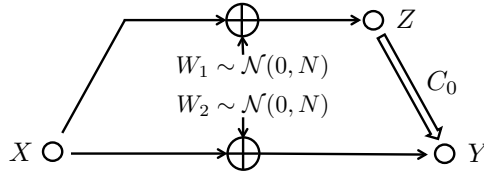


Fig. 1. Symmetric Gaussian relay channel.

For this channel it is easy to observe that

$$C(\infty) = \frac{1}{2} \log \left(1 + \frac{2P}{N} \right).$$

Let

$$C_0^* := \inf\{C_0 : C(C_0) = C(\infty)\}. \quad (2)$$

The cut-set bound [8] yields the following lower bound on C_0^* :

$$C_0^* \geq \frac{1}{2} \log \left(1 + \frac{2P}{N} \right) - \frac{1}{2} \log \left(1 + \frac{P}{N} \right),$$

which may lead one to suspect that $C(\infty)$ could be achieved at finite C_0 . The main result of our paper is to show that $C_0^* = \infty$ regardless of the parameters of the problem.

Theorem 1.1: For the symmetric Gaussian relay channel depicted in Fig. 1, $C_0^* = \infty$.

²This constraint is less stringent than requiring $\|X^n(m)\|^2 \leq nP, \forall m \in [1 : 2^{nR}]$, which is a more standard way of expressing the average power constraint for the AWGN channel. Note that the capacity can be only larger under (1) and therefore conclusions of Theorems 1.1 and 1.2 are also valid under the individual power constraint for the codewords, i.e. $\|X^n(m)\|^2 \leq nP, \forall m \in [1 : 2^{nR}]$.

This theorem follows immediately from the following theorem which establishes an upper bound on the capacity of this channel for any C_0 .

Theorem 1.2: For the symmetric Gaussian relay channel depicted in Fig. 1, the capacity $C(C_0)$ satisfies

$$\begin{cases} C(C_0) \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + C_0 + \log \sin \theta \\ C(C_0) \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + \min_{\omega \in (\frac{\pi}{2} - \theta, \frac{\pi}{2}]} h_\theta(\omega), \end{cases} \quad (3)$$

$$h_\theta(\omega) = \frac{1}{2} \log \left(\frac{4 \sin^2 \frac{\omega}{2} (P + N - N \sin^2 \frac{\omega}{2}) \sin^2 \theta}{(P + N)(\sin^2 \theta - \cos^2 \omega)} \right). \quad (4)$$

for some $\theta \in [\arcsin 2^{-C_0}, \frac{\pi}{2}]$, where

$$h_\theta(\omega) = \frac{1}{2} \log \left(\frac{4 \sin^2 \frac{\omega}{2} (P + N - N \sin^2 \frac{\omega}{2}) \sin^2 \theta}{(P + N)(\sin^2 \theta - \cos^2 \omega)} \right).$$

In Fig. 2 we plot this upper bound (label: New bound) under three different values of the SNR $= \frac{P}{N}$ of the Gaussian channels together with the celebrated cut-set bound [8] and an upper bound on the capacity of this channel we have previously derived in [6] (label: Old bound). For reference, we also provide the rate achieved by a compress-and-forward relay strategy (label: C-F), which employs Gaussian input distribution at the source combined with Gaussian quantization and Wyner-Ziv binning at the relay.³ Note that from these figures one can visually observe that the new upper bound reaches the value $C(\infty)$ only as $C_0 \rightarrow \infty$, which leads to the conclusion in Theorem 1.1. This is formally proved in the next section.

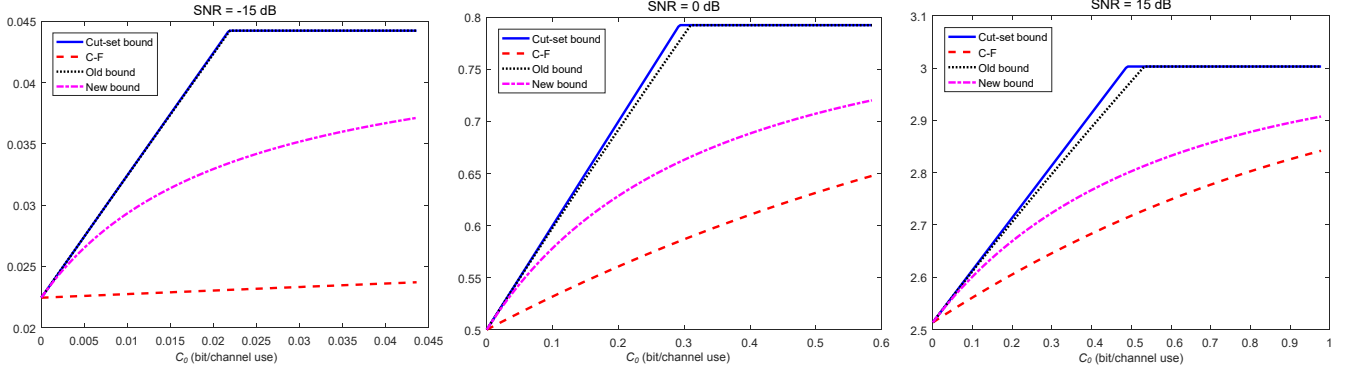


Fig. 2. Upper bounds and achievable rates for the Gaussian relay channel.

C. Approach

Our approach builds on the method we developed in our earlier work [3]–[7] for characterizing information tensions in a Markov chain by using high-dimensional geometry. The main idea is to study the geometry of the high-dimensional typical sets associated with the random variables in the Markov chain and then translate this high-dimensional geometry to information inequalities for the random variables. This idea applies equally well to single-letter and multi-letter random variables. The main geometric tool employed in our previous work [3]–[7] was the so-called blowing-up lemma. In the current paper, our main geometric ingredient is a strengthening of the isoperimetric inequality on a high-dimensional sphere, which we develop by building on the Riesz rearrangement inequality [15]. The classical isoperimetric inequality on the sphere states that among all sets on the sphere with a given volume the spherical cap has the smallest boundary or more generally the smallest volume of neighborhood [9]. In this paper, we show

³This is not the best state-of-the-art achievable rate as this rate can be improved, for example, by using bursty transmissions and time-sharing at low SNR.

that the spherical cap is the extremal set not only in terms of minimizing the volume of its neighborhood, but roughly speaking also in terms of minimizing its total intersection volume with a ball drawn around a randomly chosen point on the sphere.

It may be a priori surprising that the isoperimetric inequality appears as the main technical ingredient in the solution of a network information theory problem. However, a converse can be thought of as characterizing the extremal configuration of the typical sets of the random variables associated with an information theory problem, i.e., the configuration that is induced by the (extremal) capacity-achieving strategy. In this sense, it is quite natural that a tool, such as the isoperimetric inequality, which characterizes extremal sets in a certain geometric sense, turns out to be useful.

Formulating the problem of determining the communication capacity of channels as a problem in high-dimensional geometry is one of Shannon's most important insights that has led to the conception of information theory. His second paper [10], which appears couple of months after his classical paper "A Mathematical Theory of Communication" [11] but is cited in this first paper, develops a geometric representation of any point-to-point communication system. It then provides an elegant and intuitive geometric proof of the coding theorem for the AWGN channel, where the converse is based on a sphere-packing argument in high-dimensional space and achievability is proved by a geometric random coding argument. However, to the best of our knowledge such techniques have not been used effectively for solving network problems. Our approach is similar to Shannon's approach in [10] in that the key step in our proof is a packing argument on a spherical cap. However, it is also different from Shannon's approach as we do not directly study the geometry of the codewords but rather use high-dimensional geometry to characterize information tensions in a Markov chain by a lifting step to a high-dimensional space. We believe this approach can be useful for solving other open problems in network information theory.

II. PROOFS OF THEOREMS 1.1 AND 1.2

The proofs of both theorems follow from the below lemma, which is the main technical focus of this paper. The proof idea for this lemma is outlined in Section II-C and a formal proof is given in Section IV. We now state this lemma and show how it leads to the conclusion in Theorem 1.2, which is then used to prove Theorem 1.1.

Lemma 2.1: Let I_n be an integer random variable and X^n , Y^n and Z^n be n -length random vectors which form the Markov chain $I_n - Z^n - X^n - Y^n$. Assume moreover that Z^n and Y^n are i.i.d. white Gaussian vectors given X^n , i.e. $Z^n, Y^n \sim \mathcal{N}(X^n, N I_{n \times n})$, where $I_{n \times n}$ denotes the identity matrix, and $E[\|X^n\|^2] = nP$, and $I_n = f_n(Z^n)$ is a deterministic mapping of Z^n to a set of integers.

Let $H(I_n|X^n)$ be denoted by $-n \log \sin \theta_n$, i.e., define⁴

$$\theta_n := \arcsin 2^{-\frac{1}{n}H(I_n|X^n)}. \quad (5)$$

Then the following inequality holds for any n ,

$$H(I_n|Y^n) \leq n \cdot \min_{\omega \in (\frac{\pi}{2} - \theta_n, \frac{\pi}{2})} \frac{1}{2} \log \left(\frac{4 \sin^2 \frac{\omega}{2} (P + N - N \sin^2 \frac{\omega}{2})}{(P + N)(\sin^2 \theta_n - \cos^2 \omega)} \right). \quad (6)$$

Note that the lemma provides an upper bound on $H(I_n|Y^n)$ in terms of $H(I_n|X^n)$ for a Markov chain that satisfies the conditions of the lemma.

A. Proof of Theorem 1.2

Suppose a rate R is achievable. Then there exists a sequence of $(2^{nR}, n)$ codes such that the average probability of error $P_e^{(n)} \rightarrow 0$ as $n \rightarrow \infty$. Let the relay's transmission be denoted by $I_n = f_n(Z^n)$. By

⁴Note that values of θ_n between 0 and $\pi/2$ span all possible values for $H(I_n|X^n)$ between 0 and ∞ .

standard information theoretic arguments, for this sequence of codes we have

$$\begin{aligned} nR &= H(M) \\ &= I(M; Y^n, I_n) + H(M|Y^n, I_n) \\ &\leq I(X^n; Y^n, I_n) + n\mu \end{aligned} \quad (7)$$

$$\begin{aligned} &= I(X^n; Y^n) + I(X^n; I_n|Y^n) + n\mu \\ &= I(X^n; Y^n) + H(I_n|Y^n) - H(I_n|X^n) + n\mu \end{aligned} \quad (8)$$

$$\leq nI(X_Q; Y_Q) + H(I_n|Y^n) - H(I_n|X^n) + n\mu, \quad (9)$$

$$\leq \frac{n}{2} \log \left(1 + \frac{P}{N} \right) + H(I_n|Y^n) - H(I_n|X^n) + n\mu, \quad (10)$$

for any $\mu > 0$ and n sufficiently large. In the above, (7) follows from applying the data processing inequality to the Markov chain $M - X^n - (Y^n, I_n)$ and Fano's inequality, (8) uses the fact that $I_n - X^n - Y^n$ form a Markov chain and thus $H(I_n|X^n, Y^n) = H(I_n|X^n)$, (9) follows by defining the time sharing random variable Q to be uniformly distributed over $[1 : n]$, and (10) follows because

$$E[X_Q^2] = \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \frac{1}{n} \sum_{i=1}^n X_i^2(m) = \frac{1}{n} \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \|X^n(m)\|^2 \leq P. \quad (11)$$

Given (10), the standard way to proceed would be to upper bound the first entropy term by $H(I_n|Y^n) \leq H(I_n) \leq nC_0$ and lower bound the second entropy term $H(I_n|X^n)$ simply by 0. This would lead to the so-called multiple-access bound in the well-known cut-set bound on the capacity of this channel [8]. However as we already point out in our previous works [3]–[7], this leads to a loose bound since it does not capture the inherent tension between how large the first entropy term can be and how small the second one can be. Instead, we can use Lemma 2.1 to more tightly upper bound the difference $H(I_n|Y^n) - H(I_n|X^n)$ in (10).

We start by verifying that the random variables I_n, X^n, Z^n and Y^n associated with a code of blocklength n satisfy the conditions in the lemma. It is trivial to observe that they satisfy the required Markov chain condition and Z^n and Y^n are i.i.d. Gaussian given X^n due to the channel structure. Note also that without loss of generality we can assume that the code satisfies the average power constraint in (1) with equality, i.e.,

$$E[\|X^n\|^2] = \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \|X^n(m)\|^2 = nP.$$

This is because given a $(2^{nR}, n)$ code with average probability of error $P_e^{(n)}$ and $E[\|X^n\|^2] = nP' < nP$, we can always scale up the codewords by a factor of $\sqrt{nP/nP'}$ and achieve an average probability of error smaller than or equal to $P_e^{(n)}$.⁵

Therefore, applying Lemma 2.1 to the random variables associated with a code for the relay channel, we can bound the difference of the two entropy terms in (10) and conclude that for any achievable rate R ,

$$R \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + \min_{\omega \in \left(\frac{\pi}{2} - \theta_n, \frac{\pi}{2} \right]} h_{\theta_n}(\omega) + \mu, \quad (12)$$

where $h_{\theta_n}(\omega)$ is defined as

$$h_{\theta_n}(\omega) = \frac{1}{2} \log \left(\frac{4 \sin^2 \frac{\omega}{2} (P + N - N \sin^2 \frac{\omega}{2}) \sin^2 \theta_n}{(P + N)(\sin^2 \theta_n - \cos^2 \omega)} \right), \quad (13)$$

⁵This can be done for example by adding additional independent Gaussian noise at the relay and the destination to emulate the transmission of the original codeword.

in which $\theta_n := \arcsin 2^{-\frac{1}{n}H(I_n|X^n)}$ satisfies

$$\theta_0 := \arcsin 2^{-C_0} \leq \arcsin 2^{-\frac{1}{n}H(I_n|X^n)} = \theta_n \leq \frac{\pi}{2}. \quad (14)$$

At the same time, for any achievable rate R , we also have

$$R \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + C_0 + \log \sin \theta_n + \mu, \quad (15)$$

which simply follows from (10) by upper bounding $H(I_n|Y^n)$ with nC_0 and plugging in the definition of θ_n . Therefore, if a rate R is achievable then for any $\mu > 0$ and n sufficiently large it should simultaneously satisfy both (12) and (15) for some θ_n that satisfies the condition in (14). This concludes the proof of the theorem.

B. Proof of Theorem 1.1

In order to prove that Theorem 1.1 follows from Theorem 1.2, consider the second bound (4) on $C(C_0)$ in Theorem 1.2. Since the θ in (4) satisfies $\theta \geq \arcsin 2^{-C_0} := \theta_0$, we can upper bound the right-hand side of (4) to obtain

$$C(C_0) \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + \min_{\omega \in (\frac{\pi}{2} - \theta_0, \frac{\pi}{2}]} h_{\theta_n}(\omega).$$

Also because for any $\omega \in (\frac{\pi}{2} - \theta_0, \frac{\pi}{2}]$, $h_{\theta_n}(\omega) \leq h_{\theta_0}(\omega)$, we further have

$$C(C_0) \leq \frac{1}{2} \log \left(1 + \frac{P}{N} \right) + \min_{\omega \in (\frac{\pi}{2} - \theta_0, \frac{\pi}{2}]} h_{\theta_0}(\omega). \quad (16)$$

The significance of the function $h_{\theta_0}(\omega)$ is that for any $\theta_0 > 0$,

$$h_{\theta_0} \left(\frac{\pi}{2} \right) = \frac{1}{2} \log \left(\frac{2P + N}{P + N} \right), \quad (17)$$

and $h_{\theta_0}(\omega)$ is increasing at $\frac{\pi}{2}$, or more precisely,

$$h'_{\theta_0} \left(\frac{\pi}{2} \right) = \frac{P}{(2P + N) \ln 2} > 0.$$

Therefore, as long as $\theta_0 > 0$, which is the case when C_0 is finite, the minimization of $h_{\theta_0}(\omega)$ with respect to ω in (16) yields a value strictly smaller than $h_{\theta_0}(\frac{\pi}{2})$ in (17). This would allow us to conclude that the capacity $C(C_0)$ for any finite C_0 is strictly smaller than $\frac{1}{2} \log \left(1 + \frac{2P}{N} \right)$.

We now formalize the above argument. Using the definition of the derivative, one obtains

$$h'_{\theta_0} \left(\frac{\pi}{2} \right) = \lim_{\Delta \rightarrow 0} \frac{h_{\theta_0} \left(\frac{\pi}{2} \right) - h_{\theta_0} \left(\frac{\pi}{2} - \Delta \right)}{\Delta}.$$

Therefore, there exists a sufficiently small $\Delta_1 > 0$ such that $0 < \Delta_1 < \theta_0$ and

$$\left| \frac{h_{\theta_0} \left(\frac{\pi}{2} \right) - h_{\theta_0} \left(\frac{\pi}{2} - \Delta_1 \right)}{\Delta_1} - h'_{\theta_0} \left(\frac{\pi}{2} \right) \right| \leq \frac{h'_{\theta_0} \left(\frac{\pi}{2} \right)}{2}.$$

For such Δ_1 we have

$$\begin{aligned} h_{\theta_0} \left(\frac{\pi}{2} - \Delta_1 \right) &\leq h_{\theta_0} \left(\frac{\pi}{2} \right) - \frac{\Delta_1 h'_{\theta_0} \left(\frac{\pi}{2} \right)}{2} \\ &= \frac{1}{2} \log \left(\frac{2P + N}{P + N} \right) - \frac{P \Delta_1}{2(2P + N) \ln 2}, \end{aligned}$$

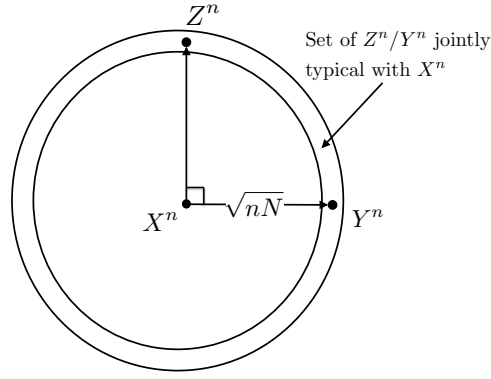


Fig. 3. Jointly typical set with X^n .

which further implies that

$$\min_{\omega \in (\frac{\pi}{2} - \theta_0, \frac{\pi}{2}]} h_{\theta_0}(\omega) \leq \frac{1}{2} \log \left(\frac{2P + N}{P + N} \right) - \frac{P\Delta_1}{2(2P + N) \ln 2}. \quad (18)$$

Combining (16) and (18) we obtain that for any finite C_0 , there exists some $\Delta_1 > 0$ such that

$$C(C_0) \leq \frac{1}{2} \log \left(1 + \frac{2P}{N} \right) - \frac{P\Delta_1}{2(2P + N) \ln 2}.$$

This proves Theorem 1.1.

C. Proof Outline for Lemma 2.1

Recall that Lemma 2.1 bounds $H(I_n|Y^n)$ in terms of $H(I_n|X^n)$ in a Markov chain $I_n - Z^n - X^n - Y^n$, where Z^n and Y^n are i.i.d. Gaussian vectors given X^n , $E[\|X^n\|^2] = nP$ and $I_n = f_n(Z^n)$ is a deterministic mapping of Z^n to a set of integers. As a preliminary exercise to check that fixing $H(I_n|X^n)$ indeed induces an upper bound on $H(I_n|Y^n)$, one can verify that if $H(I_n|X^n) = 0$ then $H(I_n|Y^n) = 0$. One (heuristic) way to see this (that can be made precise) is as follows: $H(I_n|X^n) = 0$ implies that given the transmitted codeword X^n , there is no ambiguity about I_n , or equivalently all Z^n sequences jointly typical with X^n are mapped to the same I_n . See Figure 3. However, since Y^n and Z^n are statistically equivalent given X^n (they share the same typical set given X^n) this would imply that I_n can be also determined based on Y^n and therefore $H(I_n|Y^n) = 0$.

Following a similar line of thought, if $H(I_n|X^n)$ is fixed to a certain non-zero value, say $H(I_n|X^n) = -n \log \sin \theta_n$, this roughly speaking implies that the typical Z^n 's surrounding an X^n are now mapped to multiple I_n values. This argument can be made precise as follows: Consider the following B -length i.i.d. sequence

$$\{(X^n(b), Y^n(b), Z^n(b), I_n(b))\}_{b=1}^B, \quad (19)$$

where for any $b \in [1 : B]$, $(X^n(b), Y^n(b), Z^n(b), I_n(b))$ has the same distribution as (X^n, Y^n, Z^n, I_n) . For notational convenience, in the sequel we write the B -length sequence $[X^n(1), X^n(2), \dots, X^n(B)]$ as \mathbf{X} and similarly define \mathbf{Y}, \mathbf{Z} and \mathbf{I} ; note that here we have

$$\mathbf{I} = [f_n(Z^n(1)), f_n(Z^n(2)), \dots, f_n(Z^n(B))] =: f(\mathbf{Z}).$$

Now we can apply a standard typicality argument to say that for any typical (\mathbf{x}, \mathbf{i}) pair,⁶

$$p(\mathbf{i}|\mathbf{x}) = \Pr(f(\mathbf{Z}) = \mathbf{i}|\mathbf{x}) \doteq 2^{nB \log \sin \theta_n}. \quad (20)$$

⁶Following [12], we say $a_m \doteq b_m$ if $\lim_{m \rightarrow 0} \frac{1}{m} \log \frac{a_m}{b_m} = 0$. Notations " \doteq " and " \doteq " are similarly defined.

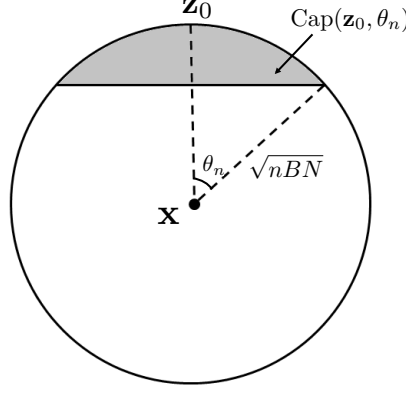


Fig. 4. A spherical cap with angle θ_n .

This probabilistic statement can be translated into the following geometric picture: Given \mathbf{x} , typical \mathbf{y} and \mathbf{z} sequences will be approximately uniformly distributed on an ϵ -thin spherical shell centered at \mathbf{x} and of radius \sqrt{nBN} , denoted as

$$\text{Shell}(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}) := \left\{ \mathbf{a} \in \mathbb{R}^{nB} : \|\mathbf{a} - \mathbf{x}\| \in \left[\sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)} \right] \right\}$$

where $\epsilon \rightarrow 0$ as $B \rightarrow \infty$. The relation (20) can then be used to argue that the set of \mathbf{z} 's jointly typical with \mathbf{x} that are mapped to the given \mathbf{i} , denoted by

$$A_{\mathbf{x}}(\mathbf{i}) = \left\{ \mathbf{z} \in \text{Shell}(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}) : f(\mathbf{z}) = \mathbf{i} \right\},$$

will occupy a volume

$$|A_{\mathbf{x}}(\mathbf{i})| \doteq 2^{nB \left(\frac{1}{2} \log 2\pi e N \sin^2 \theta_n \right)}, \quad (21)$$

on this thin shell. This translation between probabilities and volumes of sets is immediate since \mathbf{y} and \mathbf{z} are distributed approximately uniformly on the shell.

Assume now that the set $A_{\mathbf{x}}(\mathbf{i})$ were a spherical cap as illustrated in Fig. 4. In general, a spherical cap on $\text{Shell}(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)})$ can be defined as a ball in terms of the geodesic metric, or simply the angle:

$$\angle(\mathbf{y}, \mathbf{z}) = \arccos \left(\frac{\mathbf{y} \cdot \mathbf{z}}{\|\mathbf{y}\| \|\mathbf{z}\|} \right)$$

on the shell, i.e.,

$$\text{Cap}(\mathbf{z}_0, \phi) = \left\{ \mathbf{z} \in \text{Shell}(\mathbf{x}, \sqrt{nB(N-\epsilon)}, \sqrt{nB(N+\epsilon)}) : \angle(\mathbf{z}_0, \mathbf{z}) \leq \phi \right\},$$

where we will refer to \mathbf{z}_0 as the pole and ϕ as the angle of the cap. Using the volume formula for the hyperspherical cap and characterizing the exponent of such a volume (c.f. Appendix A), it can be shown that the volume in (21) would correspond to an angle of θ_n for the spherical cap as the thickness of the shell ϵ tends to zero. Now, a straightforward computation would yield the following result: Let $V_n = |\text{Cap}(\mathbf{z}_0, \theta_n) \cap \text{Cap}(\mathbf{y}_0, \omega_n)|$ where $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and $\theta_n + \omega_n > \pi/2$. Then,

$$\Pr \left(|A_{\mathbf{x}}(\mathbf{i}) \cap \text{Cap}(\mathbf{Y}, \omega_n)| \geq V_n \mid \mathbf{x} \right) \rightarrow 1 \text{ as } B \rightarrow \infty. \quad (22)$$

In words, if we take a \mathbf{y} uniformly at random on the shell and draw a spherical cap centered at \mathbf{y} with angle $\omega_n > \pi/2 - \theta_n$, then with high probability the intersection volume of this cap with the cap $A_{\mathbf{x}}(\mathbf{i})$ will be approximately lower bounded by V_n . This statement follows from the (unthinkable in low-dimensions)

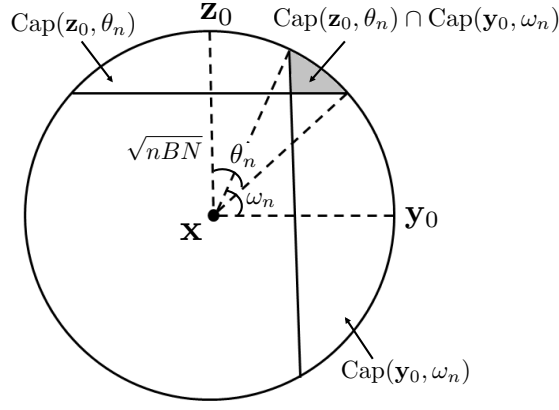


Fig. 5. Intersection of two spherical caps.

fact that in high dimensions most of the volume of the shell is concentrated around the equator (any equator), and in particular the equator at angle $\pi/2$ from the pole of $A_{\mathbf{x}}(\mathbf{i})$. Therefore, as the dimension nB gets large, for almost all \mathbf{y} 's, the intersection volume of the two spherical caps will be approximately given by V_n (see Fig. 5), which can be shown to be

$$V_n \doteq 2^{nB} \left(\frac{1}{2} \log 2\pi e N (\sin^2 \theta_n - \cos^2 \omega_n) \right),$$

by using the volume formula for the intersection of two hyperspherical caps and characterizing the exponent of this volume (c.f. Appendix B).⁷

One of the main technical steps in our proof is to show that the statement (22) holds for any arbitrary set $A_{\mathbf{x}}(\mathbf{i})$ with volume given in (21), not only when $A_{\mathbf{x}}(\mathbf{i})$ is a spherical cap as we assumed above. Note that this can be regarded as an extension of the classical isoperimetric inequality on the sphere, which states that among all sets on the sphere with a given volume, the spherical cap has the smallest boundary, or more generally the smallest volume of neighborhood. Another way to interpret the classical isoperimetric inequality is the following: given an arbitrary set A on the sphere, if we take a random point on the same sphere and draw a ball of certain radius around it, the probability that this ball touches A is at least as large as the same probability when A is a spherical cap of the same volume. Proving that (22) holds for any set amounts to saying that if we take a random point on the sphere and draw a ball of given radius, with high probability the intersection of the ball with the set A would be at least as large as the intersection we would get if A were a spherical cap. Roughly speaking, it identifies the spherical cap as the extremal set, not only for minimizing the volume of its neighborhood as done by the classical isoperimetric inequality, but also the extremal set when one is interested in minimizing the total intersection volume with A at given distance. We provide a more detailed discussion of this technical step in Section III.

The above statement allows us to reach the following conclusion regarding the random vectors $(\mathbf{I}, \mathbf{X}, \mathbf{Y})$ with high probability: if we take \mathbf{Y} and draw a Euclidean ball of radius

$$\sqrt{nBN 4 \sin^2 \frac{\omega_n}{2}} \quad (23)$$

around it, since a Euclidean ball of this radius includes the spherical cap of angle ω_n in (22); see Fig. 6, the volume of the intersection of the set $A(\mathbf{I})$ with this ball is lower bounded by

$$\left| A(\mathbf{I}) \cap \text{Ball} \left(\mathbf{Y}, \sqrt{nBN 4 \sin^2 \frac{\omega_n}{2}} \right) \right| \geq 2^{nB} \left(\frac{1}{2} \log 2\pi e N (\sin^2 \theta_n - \cos^2 \omega_n) \right), \quad (24)$$

⁷Precisely speaking, Appendix B provides the intersection area of two hyperspherical caps on a sphere rather than the intersection volume of two caps on a shell. Note however that as the dimension grows the intersection volume and area, computed on the shell and on the sphere respectively, approach the same in the exponent.

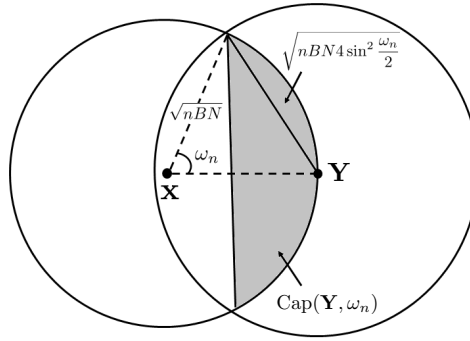
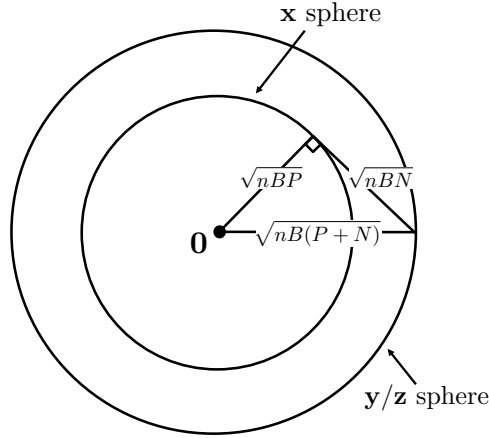


Fig. 6. Euclidean ball contains the cap.

Fig. 7. \mathbf{x} sphere and \mathbf{y}/\mathbf{z} sphere.

where $A(\mathbf{I})$ is defined as $A(\mathbf{I}) = \{\mathbf{z} \in \mathbb{R}^{nB} : f(\mathbf{z}) = \mathbf{I}\}$ and $\text{Ball}(\mathbf{c}, r)$ denotes a ball centered at \mathbf{c} with radius r . This follows from (22), since (22) says that this property holds with high probability conditioned on any \mathbf{x} which is typical with (\mathbf{I}, \mathbf{Y}) . In words, if we take a typical realization (\mathbf{i}, \mathbf{y}) of (\mathbf{I}, \mathbf{Y}) and draw a ball of radius (23) around \mathbf{y} , the volume of the set of points that are mapped to \mathbf{i} in this ball is lower bounded as in (24). This puts an upper limit on the number of possible values of \mathbf{i} given \mathbf{y} . To get a tighter bound, we can incorporate the fact that most of the \mathbf{x} 's lie on a thin shell of radius \sqrt{nBP} , and \mathbf{y} and \mathbf{z} lie on a thin shell of radius $\sqrt{nB(P+N)}$. See Fig. 7. Therefore the number of possible values for \mathbf{I} given \mathbf{Y} can be bounded by the ratio of the spherical cap volume

$$\left| \text{Shell}\left(\mathbf{0}, \sqrt{nB(P+N-\epsilon)}, \sqrt{nB(P+N+\epsilon)}\right) \cap \text{Ball}\left(\sqrt{nB(P+N)}\mathbf{e}, \sqrt{nBN4\sin^2\frac{\omega_n}{2}}\right) \right|,$$

where \mathbf{e} is any arbitrary unit vector, to the volume each possible \mathbf{i} occupies from this cap

$$2^{nB\left(\frac{1}{2} \log 2\pi eN(\sin^2\theta_n - \cos^2\omega_n)\right)}.$$

See Figure 8. This ratio can be shown to be

$$\leq 2^{nB\left[\frac{1}{2} \log \frac{4\sin^2\frac{\omega_n}{2}(P+N-N\sin^2\frac{\omega_n}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega_n)}\right]},$$

which in turn imposes the following bound on $H(I_n|Y^n)$:

$$H(I_n|Y^n) \leq n \left(\frac{1}{2} \log \frac{4\sin^2\frac{\omega_n}{2}(P+N-N\sin^2\frac{\omega_n}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega_n)} \right).$$

The upper bound (6) in Lemma 2.1 follows by noting that the above argument holds for any $\omega_n > \pi/2 - \theta_n$.

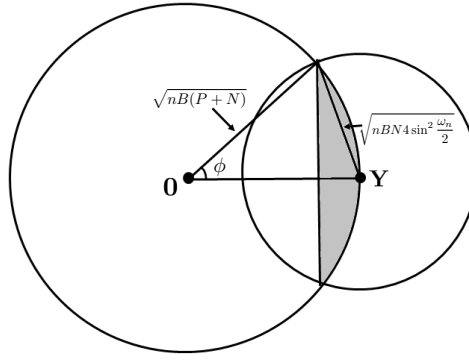


Fig. 8. A spherical cap with angle $\phi = 2 \arcsin \sqrt{\frac{N \sin^2 \frac{\omega_n}{2}}{P+N}}$.

III. STRENGTHENING THE ISOPERIMETRIC INEQUALITY

We now provide a more detailed discussion of the main geometric tool we develop for proving Lemma 2.1.

As mentioned in the previous section, a key ingredient of our result is a strengthening of the isoperimetric inequality on the sphere, which we obtain by building on the Riesz rearrangement inequality. The classical isoperimetric inequality on the sphere states that among all sets on the sphere with a given volume, the spherical cap has the smallest boundary, or more generally the smallest volume of neighborhood. This can be formally stated as follows: Let $\mathbb{S}^{m-1} \subseteq \mathbb{R}^m$ denote the $(m-1)$ -sphere of radius R , i.e. $\mathbb{S}^{m-1} = \{\mathbf{z} \in \mathbb{R}^m : \|\mathbf{z}\| = R\}$, equipped with the rotation invariant (Haar) measure $\mu = \mu_{m-1}$ normalized such that

$$\mu(\mathbb{S}^{m-1}) = \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} R^{m-1},$$

i.e. the usual surface area, and let $\mathbb{P}(A)$ denote the probability of a set or event A with respect to the corresponding Haar probability measure, i.e. normalized Haar measure such that $\mathbb{P}(\mathbb{S}^{m-1}) = 1$. A spherical cap can be defined as a ball in the metric $\angle(\mathbf{z}, \mathbf{y}) = \arccos(\langle \mathbf{z}/R, \mathbf{y}/R \rangle)$ of \mathbb{S}^{m-1} , i.e.,

$$\text{Cap}(\mathbf{z}_0, \theta) = \{\mathbf{z} \in \mathbb{S}^{m-1} : \angle(\mathbf{z}_0, \mathbf{z}) \leq \theta\}.$$

The isoperimetric inequality states that for any arbitrary set $A \subseteq \mathbb{S}^{m-1}$ such that $\mu(A) = \mu(C)$, where $C = \text{Cap}(\mathbf{z}_0, \theta) \subseteq \mathbb{S}^{m-1}$ is a spherical cap, it holds that

$$\mu(A_t) \geq \mu(C_t), \quad \forall t \geq 0,$$

where A_t is the t -neighborhood of A , defined as

$$A_t = \left\{ \mathbf{z} \in \mathbb{S}^{m-1} : \min_{\mathbf{z}' \in A} \angle(\mathbf{z}, \mathbf{z}') \leq t \right\},$$

and similarly

$$C_t = \left\{ \mathbf{z} \in \mathbb{S}^{m-1} : \min_{\mathbf{z}' \in C} \angle(\mathbf{z}, \mathbf{z}') \leq t \right\} = \text{Cap}(\mathbf{z}_0, \theta + t).$$

The isoperimetric inequality above, combined with a standard computation, immediately yields the following measure concentration result: Let $A \subseteq \mathbb{S}^{m-1}$ be an arbitrary set and $C = \text{Cap}(\mathbf{z}_0, \theta) \subseteq \mathbb{S}^{m-1}$ be a spherical cap such that $\mu(A) = \mu(C)$. If this is the case, we will say that A has an effective angle of θ . Then for any $\epsilon > 0$ and m sufficiently large,

$$\mathbb{P}(A_{\frac{\pi}{2} - \theta + \epsilon}) \geq 1 - \epsilon. \quad (25)$$

This is known as the blowing-up lemma on the sphere. Note that when A is a spherical cap, this result follows from an elementary computation which reveals that most of the volume of an $(m-1)$ -sphere is concentrated around the boundary of any half-sphere, i.e. any equator, therefore blowing-up the spherical cap to be slightly bigger than a half-sphere is sufficient to capture most of the volume or probability mass on the sphere. The isoperimetric inequality allows to extend this statement to any arbitrary set A and make the much more general statement that the volume is concentrated around the boundary of any set with probability $1/2$.

An equivalent way to view this result is the following: let $A \subseteq \mathbb{S}^{m-1}$ be an arbitrary set with effective angle $\theta > 0$, then

$$\mathbb{P}\left(\mu\left(A \cap \text{Cap}\left(\mathbf{Y}, \frac{\pi}{2} - \theta + \epsilon\right)\right) > 0\right) > 1 - \epsilon, \quad (26)$$

for any $\epsilon > 0$ and sufficiently large m . In words, if we take a \mathbf{y} uniformly at random on the sphere and draw a spherical cap of angle slightly larger than $\frac{\pi}{2} - \theta$ around it (or equivalently a Euclidean ball centered at \mathbf{y} of radius equal to the corresponding Euclidean distance), this cap will intersect the set A with high probability. This statement is almost equivalent to (25) since the \mathbf{y} 's for which the intersection is non-zero lie in the $\frac{\pi}{2} - \theta + \epsilon$ -neighborhood of A . Note again that this statement would be trivial if A were known to be a spherical cap, and it holds for any A due to the isoperimetric inequality.

By building on the Riesz rearrangement inequality in the next subsection, we show the following stronger result:

Lemma 3.1: Let $A \subseteq \mathbb{S}^{m-1}$ be any arbitrary set with effective angle $\theta > 0$ (i.e. $\mu(A) = \mu(C)$ with $C = \text{Cap}(\mathbf{z}_0, \theta)$ for some $\mathbf{z}_0 \in \mathbb{S}^{m-1}$), and let $V = \mu(\text{Cap}(\mathbf{z}_0, \theta) \cap \text{Cap}(\mathbf{y}_0, \omega))$ for some $\mathbf{z}_0 \in \mathbb{S}^{m-1}$ and $\mathbf{y}_0 \in \mathbb{S}^{m-1}$ such that $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and $\theta + \omega > \pi/2$. Then for any $\epsilon > 0$ and m sufficiently large,

$$\mathbb{P}(\mu(A \cap \text{Cap}(\mathbf{Y}, \omega + \epsilon)) > V) \geq 1 - \epsilon.$$

Note that if A itself is a cap the statement is straightforward and follows from the fact that \mathbf{y} with high probability will be concentrated around the equator at angle $\pi/2$ from the pole of A . Therefore, as m gets large for almost all \mathbf{y} 's, the intersection of the two spherical caps will be given by V . The statement however is much stronger than this and holds for any arbitrary set A , analogous to (26). It states that no matter what the set A is, if we take a random point on the sphere and draw a cap of angle slightly larger than ω , then with high probability the intersection of the cap with the set A would be at least as large as the intersection we would get if A were a spherical cap. Roughly speaking, it identifies the spherical cap as the extremal set, not only for minimizing the volume of its t -neighborhood as done by the isoperimetric inequality, i.e. when one is interested in touching a single point in A , but also the extremal set when one is interested in minimizing the total intersection volume with A at certain distance.

As we saw in the previous section, what we actually need to prove Lemma 2.1 is a version of Lemma 3.1 on a thin spherical shell instead of on the sphere, since the typical \mathbf{z} and \mathbf{y} sequences in our problem concentrate on a thin shell, which can be made thinner and thinner as the dimension increases, but the sphere itself has measure zero compared to a shell. However, it is instructive to first prove Lemma 3.1, which can be interesting in its own right. We will discuss the extension to a spherical shell in Section V-B.

A. Proof of Lemma 3.1 via Spherical Rearrangements

We prove Lemma 3.1 via spherical rearrangements in two steps. We first verify below that the claim in Lemma 3.1 holds when A is known to be a spherical cap (with the intersection volume actually larger than $2V$ instead of just V). We then show that if it holds for a spherical cap, then it holds for any arbitrary set A .

Assume $A = \text{Cap}(\mathbf{z}_0, \theta)$ for some $\mathbf{z}_0 \in \mathbb{S}^{m-1}$. Due to measure concentration, for any $\epsilon > 0$ and m sufficiently large, we have

$$\mathbb{P}(\angle(\mathbf{z}_0, \mathbf{Y}) \in [\pi/2 - \epsilon/2, \pi/2 + \epsilon/2]) \geq 1 - \epsilon/2.$$

Due to the triangle inequality for the geodesic metric (angle), for \mathbf{y} such that $\angle(\mathbf{z}_0, \mathbf{y}) \in [\pi/2 - \epsilon/2, \pi/2 + \epsilon/2]$ we have

$$A \cap \text{Cap}(\mathbf{y}_0, \omega) \subseteq A \cap \text{Cap}(\mathbf{y}, \omega + \epsilon/2)$$

where \mathbf{y}_0 is such that $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$. Therefore,

$$\mathbb{P}(\mu(A \cap \text{Cap}(\mathbf{Y}, \omega + \epsilon/2)) > V) \geq 1 - \epsilon/2.$$

Indeed, because the intersection volume is growing exponentially in the angle $\omega + \epsilon/2$, we can also conclude that for any $\epsilon > 0$ and m sufficiently large,

$$\mathbb{P}(\mu(A \cap \text{Cap}(\mathbf{Y}, \omega + \epsilon)) > 2V) \geq 1 - \epsilon/2. \quad (27)$$

Our main tool for extending the above result to arbitrary A is the symmetric decreasing rearrangement of functions on the sphere, along with a version of the Riesz rearrangement inequality on the sphere due to Baernstein and Taylor [15].

For any function $f : \mathbb{S}^{m-1} \rightarrow \mathbb{R}$ and pole \mathbf{z}_0 , the symmetric decreasing rearrangement of f about \mathbf{z}_0 is defined to be the function $f^* : \mathbb{S}^{m-1} \rightarrow \mathbb{R}$ such that $f^*(\mathbf{y})$ depends only on the angle $\angle(\mathbf{y}, \mathbf{z}_0)$, is nonincreasing in $\angle(\mathbf{y}, \mathbf{z}_0)$, and has super-level sets of the same size as f , i.e.

$$\mu(\{\mathbf{y} : f^*(\mathbf{y}) > d\}) = \mu(\{\mathbf{y} : f(\mathbf{y}) > d\})$$

for all d . The function f^* is unique except for on sets of measure zero.

One important special case is when the function $f = 1_A$ is the characteristic function for a subset A . The function 1_A is just the function such that

$$1_A(\mathbf{y}) = \begin{cases} 1 & \mathbf{y} \in A \\ 0 & \text{otherwise.} \end{cases}$$

In this case 1_A^* is equal to the characteristic function associated with a spherical cap of the same size as A . In other words, if A^* is a spherical cap about the pole \mathbf{z}_0 such that $\mu(A^*) = \mu(A)$, then $1_A^* = 1_{A^*}$.

Lemma 3.2 (Baernstein and Taylor [15]): Let K be a nondecreasing bounded measurable function on the interval $[-1, 1]$. Then for all functions $f, g \in L^1(\mathbb{S}^{m-1})$,

$$\begin{aligned} & \int_{\mathbb{S}^{m-1}} \left(\int_{\mathbb{S}^{m-1}} f(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} \right) g(\mathbf{y}) d\mathbf{y} \\ & \leq \int_{\mathbb{S}^{m-1}} \left(\int_{\mathbb{S}^{m-1}} f^*(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} \right) g^*(\mathbf{y}) d\mathbf{y}. \end{aligned}$$

In order to prove Lemma 3.1, we will apply Lemma 3.2 by choosing f, K such that the inner integral

$$\int_{\mathbb{S}^{m-1}} f(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} = \mu(A \cap \text{Cap}(\mathbf{y}, \omega + \epsilon)).$$

To do this, given an arbitrary set A , we set $f = 1_A$ and

$$K(\cos \alpha) = \begin{cases} 1 & 0 \leq \alpha \leq \omega + \epsilon \\ 0 & \omega + \epsilon < \alpha \leq \pi. \end{cases}$$

Note that K is nondecreasing, bounded, and measurable. Furthermore, the product $f(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle)$ is one precisely when $\mathbf{z} \in A$ and $\angle(\mathbf{z}, \mathbf{y}) \leq \omega + \epsilon$, and it is zero otherwise. Thus the integral

$$\psi(\mathbf{y}) = \int_{\mathbb{S}^{m-1}} f(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} \quad (28)$$

is exactly the measure of the set $A \cap \text{Cap}(\mathbf{y}, \omega + \epsilon)$. Recall that the spherical cap $\text{Cap}(\mathbf{y}, \omega + \epsilon)$ is centered around the pole \mathbf{y} with angle $\omega + \epsilon$.

We will use test functions g which are also characteristic functions. Let $g = 1_C$ for some measurable subset $C \subseteq \mathbb{S}^{m-1}$. For a fixed measure $\mu(C)$, the left-hand side of the inequality from Lemma 3.2 will be maximized by picking C to correspond to a super-level set of ψ (this can be shown easily by contradiction, although we don't actually use this fact and it just provides the motivation).

The claim is that when using $g = 1_C$ where $C = \{\mathbf{y} : \psi(\mathbf{y}) > d\}$ for some d , we have the following equality:

$$\begin{aligned} \int_{\mathbb{S}^{m-1}} \psi(\mathbf{y}) 1_C(\mathbf{y}) d\mathbf{y} &= \int_{\mathbb{S}^{m-1}} \psi^*(\mathbf{y}) 1_C^*(\mathbf{y}) d\mathbf{y} \\ &= \int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y}. \end{aligned}$$

This follows from the layer-cake decomposition in that

$$\begin{aligned} \int_{\mathbb{S}^{m-1}} \psi(\mathbf{y}) 1_C(\mathbf{y}) d\mathbf{y} &= \int_C \psi(\mathbf{y}) d\mathbf{y} \\ &= \int_C \int_0^\infty 1_{\{\psi(\mathbf{y}) > t\}} dt d\mathbf{y} \\ &= \int_0^\infty \int_C 1_{\{\psi(\mathbf{y}) > t\}} d\mathbf{y} dt \\ &= \int_0^d \int_C 1_{\{\psi(\mathbf{y}) > t\}} d\mathbf{y} dt + \int_d^\infty \int_C 1_{\{\psi(\mathbf{y}) > t\}} d\mathbf{y} dt \\ &= \int_0^d \int_{C^*} 1_{\{\psi^*(\mathbf{y}) > t\}} d\mathbf{y} dt + \int_d^\infty \int_{C^*} 1_{\{\psi^*(\mathbf{y}) > t\}} d\mathbf{y} dt \\ &= \int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y}. \end{aligned} \tag{29}$$

Using this equality and our choices for g, f, K above we will rewrite the inequality from Lemma 3.2 as

$$\int_{C^*} \psi^*(\mathbf{y}) d\mathbf{y} \leq \int_{C^*} \psi'(\mathbf{y}) d\mathbf{y} \tag{30}$$

where

$$\psi'(\mathbf{y}) = \int_{\mathbb{S}^{m-1}} f^*(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} = \mu(A^* \cap \text{Cap}(\mathbf{y}, \omega + \epsilon)).$$

Note that both $\psi^*(\mathbf{y})$ and $\psi'(\mathbf{y})$ are spherically symmetric and nonincreasing. More concretely, they both depend only on the angle $\angle(\mathbf{y}, \mathbf{z}_0)$ and are nonincreasing in the angle $\angle(\mathbf{y}, \mathbf{z}_0)$. In an abuse of notation we will write $\psi'(\alpha)$ and $\psi^*(\alpha)$ where $\alpha = \angle(\mathbf{y}, \mathbf{z}_0)$. For convenience we will define a measure ν by

$$d\nu(\phi) = A_{m-2}(R \sin \phi) R d\phi$$

where $A_m(R)$ denotes the Haar measure of the m -sphere with radius R . We do this so that an integral like

$$\int_{\mathbb{S}^{m-1}} \psi^* d\mathbf{y} = \int_0^\pi \psi^*(\phi) A_{m-2}(R \sin \phi) R d\phi$$

can be expressed as

$$\int_0^\pi \psi^* d\nu.$$

We are now ready to prove Lemma 3.1. Define ξ to be such that

$$\psi' \left(\frac{\pi}{2} + \xi \right) = 2V .$$

It is easy to see that ξ is uniquely defined and will be greater than zero for sufficiently large m . We also get from (27) that

$$\mathbb{P}(\psi'(\mathbf{Y}) > 2V) = \frac{1}{A_{m-1}(R)} \int_0^{\frac{\pi}{2} + \xi} d\nu \geq 1 - \epsilon/2 \quad (31)$$

for any $\epsilon > 0$ and m sufficiently large.

Suppose there is a subsequence m_k such that for each $m = m_k$, there exists a measurable subset A with $\psi^*(\beta) = V$ for a $0 < \beta < \pi/2 + \xi$. More specifically, let β be the smallest value such that $\psi^*(\beta) = V$. This quantity is well-defined, i.e. the minimum is achieved, because of the continuity of ψ^* . Our goal will be to show that even for such a subsequence,

$$\mathbb{P}(\psi'(\mathbf{Y}) > 2V \text{ and } \psi^*(\mathbf{Y}) \leq V) = \frac{1}{A_{m-1}(R)} \int_{\beta}^{\frac{\pi}{2} + \xi} d\nu \leq \epsilon/2$$

as $m = m_k$ increases. Together with (31), this would imply that

$$\mathbb{P}(\psi^*(\mathbf{Y}) \leq V) \leq \epsilon,$$

and because we have

$$\mathbb{P}(\psi^*(\mathbf{Y}) > V) = \mathbb{P}(\psi(\mathbf{Y}) > V),$$

we would have

$$\mathbb{P}(\mu(A \cap \mathbf{Cap}(\mathbf{Y}, \omega + \epsilon)) > V) \geq 1 - \epsilon,$$

for any arbitrary set A .

To this end, we note that

$$\frac{1}{A_{m-1}(R)} \int_{\frac{\pi}{2} + \xi}^{\pi} d\nu \leq \epsilon/2 \quad (32)$$

as a consequence of (31). However, we have the following chain of (in)equalities which will be justified below.

$$\frac{1}{A_{m-1}(R)} \int_{\frac{\pi}{2} + \xi}^{\pi} d\nu \geq \frac{1}{VA_{m-1}(R)} \int_{\frac{\pi}{2} + \xi}^{\pi} (\psi^* - \psi') d\nu \quad (33)$$

$$= \frac{1}{VA_{m-1}(R)} \int_0^{\frac{\pi}{2} + \xi} (\psi' - \psi^*) d\nu \quad (34)$$

$$\geq \frac{1}{VA_{m-1}(R)} \int_{\beta}^{\frac{\pi}{2} + \xi} (\psi' - \psi^*) d\nu \quad (35)$$

$$\geq \frac{1}{A_{m-1}(R)} \int_{\beta}^{\frac{\pi}{2} + \xi} d\nu \quad (36)$$

From these inequalities, (36) must also be bounded by $\epsilon/2$ and we are done.

The first inequality (33) is a consequence of the fact that over the range of the integral, ψ^* is less than or equal to V and ψ' is non-negative. The equality in (34) follows from

$$\int_0^{\pi} \psi^* d\nu = \int_0^{\pi} \psi' d\nu ,$$

which is itself a consequence of (29) with $C = \mathbb{S}^{m-1}$ and

$$\begin{aligned}
& \int_{\mathbb{S}^{m-1}} \int_{\mathbb{S}^{m-1}} f(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} d\mathbf{y} \\
&= \int \int K(\langle \mathbf{y}, \mathbf{z} \rangle) d\mathbf{y} f(\mathbf{z}) d\mathbf{z} \\
&= \int \mu(\text{Cap}(\mathbf{y}, \omega + \epsilon)) f(\mathbf{z}) d\mathbf{z} \\
&= \mu(\text{Cap}(\mathbf{y}, \omega + \epsilon)) \mu(A) \\
&= \int \mu(\text{Cap}(\mathbf{y}, \omega + \epsilon)) f^*(\mathbf{z}) d\mathbf{z} \\
&= \int \int f^*(\mathbf{z}) K(\langle \mathbf{z}/R, \mathbf{y}/R \rangle) d\mathbf{z} d\mathbf{y} .
\end{aligned}$$

Next we have (35) which is due to the rearrangement inequality (30) when C^* is a spherical cap with angle β . Finally, the inequality (36) is because $\psi'(\mathbf{y}) \geq 2V$ and $\psi^* \leq V$ over the range of the integral.

IV. PROOF OF LEMMA 2.1

We are now in a position to prove Lemma 2.1. For this we will look at B -length i.i.d. sequences of the random vectors X^n, Y^n, Z^n , and I_n , and derive some typicality properties for these sequences which hold with high probability when B is large.

Specifically, consider the following B -length i.i.d. sequence

$$\{(X^n(b), Y^n(b), Z^n(b), I_n(b))\}_{b=1}^B, \quad (37)$$

where for any $b \in [1 : B]$, $(X^n(b), Y^n(b), Z^n(b), I_n(b))$ has the same distribution as (X^n, Y^n, Z^n, I_n) . For notational convenience, in the sequel we write the B -length sequence $[X^n(1), X^n(2), \dots, X^n(B)]$ as \mathbf{X} and similarly define \mathbf{Y}, \mathbf{Z} and \mathbf{I} ; note that here we have $\mathbf{I} = [f_n(Z^n(1)), f_n(Z^n(2)), \dots, f_n(Z^n(B))] =: f(\mathbf{Z})$.

We now introduce two lemmas that will be used to establish Lemma 2.1. The proof of the first lemma will be provided at the end of this section after we finish the proof of Lemma 2.1, while the proof of the second one contains significant novelty both conceptually and technically and deserves a dedicated treatment in Section V.

Lemma 4.1: Let $\text{Shell}(\mathbf{c}, r_1, r_2)$ denote a spherical shell defined as

$$\text{Shell}(\mathbf{c}, r_1, r_2) := \{\mathbf{a} \in \mathbb{R}^{nB} : r_1 \leq \|\mathbf{a} - \mathbf{c}\| \leq r_2\}.$$

For any $\delta > 0$ and B sufficiently large, we have

$$\begin{aligned}
& \Pr(E_1) \geq 1 - \delta \\
& \text{and } \Pr(E_2) \geq 1 - \delta,
\end{aligned}$$

where

$$E_1 := \left\{ \mathbf{Z} \in \text{Shell}\left(\mathbf{0}, \sqrt{nB(P + N - \delta)}, \sqrt{nB(P + N + \delta)}\right) \right\} \quad (38)$$

$$\text{and } E_2 := \left\{ \mathbf{Y} \in \text{Shell}\left(\mathbf{0}, \sqrt{nB(P + N - \delta)}, \sqrt{nB(P + N + \delta)}\right) \right\}. \quad (39)$$

Lemma 4.2: Let $\text{Ball}(\mathbf{c}, r)$ denote a Euclidean ball centered at \mathbf{c} and of radius r , i.e.,

$$\text{Ball}(\mathbf{c}, r) := \{\mathbf{a} \in \mathbb{R}^{nB} : \|\mathbf{a} - \mathbf{c}\| \leq r\}.$$

For any $\delta > 0$ and B sufficiently large, we have

$$\Pr(E_3) \geq 1 - \delta,$$

where E_3 is defined to be the following event:

$$\left\{ \left| f^{-1}(\mathbf{I}) \cap \text{Ball} \left(\mathbf{0}, \sqrt{nB(P+N+\delta)} \right) \cap \text{Ball} \left(\mathbf{Y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \delta \right)} \right) \right| \geq 2^{nB[\frac{1}{2} \log(2\pi eN(\sin^2 \theta_n - \cos^2 \omega)) - \delta]} \right\} \quad (40)$$

in which $f^{-1}(\mathbf{I}) := \{\mathbf{a} \in \mathbb{R}^{nB} : f(\mathbf{a}) = \mathbf{I}\}$, θ_n is given by $\arcsin 2^{-\frac{1}{n}H(I_n|X^n)}$ as in (5) and $\omega \in (\pi/2 - \theta_n, \pi/2]$.

With the above two lemmas, we now upper bound $H(\mathbf{I}|\mathbf{Y})$. Consider the indicator function

$$F = \mathbb{I}(E_1, E_2, E_3)$$

where $\mathbb{I}(\cdot)$ is defined as

$$\mathbb{I}(A) = \begin{cases} 1 & \text{if } A \text{ holds} \\ 0 & \text{otherwise,} \end{cases}$$

and the events E_1, E_2 and E_3 are as given by (38)–(40). Obviously, by the union bound, we have

$$\Pr(F = 1) \geq 1 - 3\delta$$

for any $\delta > 0$ and B sufficiently large, and therefore

$$\begin{aligned} H(\mathbf{I}|\mathbf{Y}) &\leq H(\mathbf{I}, F|\mathbf{Y}) \\ &= H(F|\mathbf{Y}) + H(\mathbf{I}|\mathbf{Y}, F) \\ &\leq H(\mathbf{I}|\mathbf{Y}, F) + 1 \\ &= \Pr(F = 1)H(\mathbf{I}|\mathbf{Y}, F = 1) + \Pr(F = 0)H(\mathbf{I}|\mathbf{Y}, F = 0) + 1 \\ &\leq H(\mathbf{I}|\mathbf{Y}, F = 1) + 3\delta nBC_0 + 1. \end{aligned} \quad (41)$$

To bound $H(\mathbf{I}|\mathbf{Y}, F = 1)$, we apply a packing argument as follows. Consider a ball centered at \mathbf{Y} and of radius $\sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \delta \right)}$, i.e.,

$$\text{Ball} \left(\mathbf{Y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \delta \right)} \right),$$

where

$$\mathbf{Y} \in \text{Shell} \left(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)} \right),$$

and ω satisfies

$$\pi/2 - \theta_n < \omega \leq \pi/2.$$

We now upper bound the volume of the intersection between this ball and $\text{Ball} \left(\mathbf{0}, \sqrt{nB(P+N+\delta)} \right)$, i.e.,

$$\left| \text{Ball} \left(\mathbf{Y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \delta \right)} \right) \cap \text{Ball} \left(\mathbf{0}, \sqrt{nB(P+N+\delta)} \right) \right|.$$

For this, first consider the intersection volume

$$\left| \text{Ball} \left(\mathbf{Y}, \sqrt{nBN 4\sin^2 \frac{\omega}{2}} \right) \cap \text{Ball} \left(\mathbf{0}, \sqrt{nB(P+N)} \right) \right|, \quad (42)$$

for

$$\mathbf{Y} \in \text{Sphere} \left(\mathbf{0}, \sqrt{nB(P+N)} \right),$$

where $\text{Sphere}(\mathbf{c}, r)$ denotes a sphere centered at \mathbf{c} of radius r , i.e.,

$$\text{Sphere}(\mathbf{c}, r) := \{\mathbf{a} \in \mathbb{R}^{nB} : \|\mathbf{a} - \mathbf{c}\| = r\}.$$

See Fig. 8. Note that the intersection in (42) actually consists of two spherical caps that are respectively located on $\text{Ball}(\mathbf{Y}, \sqrt{nBN4\sin^2\frac{\omega}{2}})$ and $\text{Ball}(\mathbf{0}, \sqrt{nB(P+N)})$. Since these two caps have essentially the same volume, it is sufficient to calculate the volume for the cap on $\text{Ball}(\mathbf{0}, \sqrt{nB(P+N)})$, i.e. the shaded area in Fig. 8. In particular, observe that

$$\begin{aligned}\phi &= 2 \arcsin \frac{\sqrt{nBN\sin^2\frac{\omega}{2}}}{\sqrt{nB(P+N)}} \\ &= 2 \arcsin \sqrt{\frac{N\sin^2\frac{\omega}{2}}{P+N}} \\ &= 2\phi_1\end{aligned}$$

where $\phi_1 := \arcsin \sqrt{\frac{N\sin^2\frac{\omega}{2}}{P+N}}$. Using the volume formula for the hyperspherical cap and characterizing the exponent of such a volume (c.f. Appendix A), we have that for any $\mathbf{Y} \in \text{Sphere}(\mathbf{0}, \sqrt{nB(P+N)})$ and B sufficiently large,

$$\left| \text{Ball}\left(\mathbf{Y}, \sqrt{nBN4\sin^2\frac{\omega}{2}}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N)}\right) \right| \leq 2^{nB[\frac{1}{2} \log(2\pi e(P+N)\sin^2\phi) + \delta]}.$$

Since

$$\begin{aligned}\sin^2\phi &= (2\sin\phi_1 \cos\phi_1)^2 \\ &= 4\sin^2\phi_1 \cos^2\phi_1 \\ &= 4\frac{N\sin^2\frac{\omega}{2}}{P+N} \left(1 - \frac{N\sin^2\frac{\omega}{2}}{P+N}\right),\end{aligned}$$

we have for any $\mathbf{Y} \in \text{Sphere}(\mathbf{0}, \sqrt{nB(P+N)})$ and B sufficiently large,

$$\left| \text{Ball}\left(\mathbf{Y}, \sqrt{nBN4\sin^2\frac{\omega}{2}}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N)}\right) \right| \leq 2^{nB\left[\frac{1}{2} \log \frac{8\pi e N \sin^2\frac{\omega}{2} (P+N - N\sin^2\frac{\omega}{2})}{P+N} + \delta\right]}.$$

By continuity, this would further imply that for any $\mathbf{Y} \in \text{Shell}(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)})$,

$$\left| \text{Ball}\left(\mathbf{Y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right) \right| \leq 2^{nB\left[\frac{1}{2} \log \frac{8\pi e N \sin^2\frac{\omega}{2} (P+N - N\sin^2\frac{\omega}{2})}{P+N} + \delta_1\right]},$$

where $\delta_1 \rightarrow 0$ as $\delta \rightarrow 0$ and $B \rightarrow \infty$.

On the other hand, the condition $F = 1$ (c.f. the definition of E_3 in Lemma 4.2) also ensures that given \mathbf{Y} , each possible \mathbf{I} occupies a volume of at least

$$2^{nB[\frac{1}{2} \log(2\pi e N (\sin^2\theta_n - \cos^2\omega)) - \delta]}$$

from the set

$$\text{Ball}\left(\mathbf{Y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right).$$

Therefore, given $F = 1$ and $\mathbf{Y} \in \text{Shell}(\mathbf{0}, \sqrt{nB(P+N-\delta)}, \sqrt{nB(P+N+\delta)})$, the number of different possibilities for \mathbf{I} is upper bounded by the ratio between

$$\left| \text{Ball}\left(\mathbf{Y}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \delta\right)}\right) \cap \text{Ball}\left(\mathbf{0}, \sqrt{nB(P+N+\delta)}\right) \right|$$

and

$$2^{nB[\frac{1}{2}\log(2\pi eN(\sin^2\theta_n - \cos^2\omega)) - \delta]},$$

which can be further upper bounded by

$$\begin{aligned} & 2^{nB\left[\frac{1}{2}\log\frac{8\pi eN\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{P+N} - \frac{1}{2}\log(2\pi eN(\sin^2\theta_n - \cos^2\omega)) + \delta + \delta_1\right]} \\ &= 2^{nB\left[\frac{1}{2}\log\frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega)} + \delta_2\right]}, \end{aligned}$$

where $\delta_2 \rightarrow 0$ as $\delta \rightarrow 0$ and $B \rightarrow \infty$. This immediately implies the following upper bound on $H(\mathbf{I}|\mathbf{Y}, F = 1)$,

$$H(\mathbf{I}|\mathbf{Y}, F = 1) \leq nB \left[\frac{1}{2} \log \frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega)} + \delta_2 \right],$$

which combined with (41) yields that

$$H(\mathbf{I}|\mathbf{Y}) \leq nB \left[\frac{1}{2} \log \frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega)} + \delta_2 \right] + 3\delta nBC_0 + 1.$$

Dividing B at both sides of the above inequality and noting that

$$H(\mathbf{I}|\mathbf{Y}) = \sum_{b=1}^B H(I_n(b)|Y^n(b)) = BH(I_n|Y^n),$$

we have

$$H(I_n|Y^n) \leq n \left(\frac{1}{2} \log \frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega)} + \delta_2 + 3\delta C_0 + \frac{1}{nB} \right). \quad (43)$$

Since δ, δ_2 and $\frac{1}{nB}$ in (43) can all be made arbitrarily small by choosing B sufficiently large, we obtain

$$H(I_n|Y^n) \leq n \left(\frac{1}{2} \log \frac{4\sin^2\frac{\omega}{2}(P+N-N\sin^2\frac{\omega}{2})}{(P+N)(\sin^2\theta_n - \cos^2\omega)} \right). \quad (44)$$

Noting that (44) holds for any $\omega \in (\frac{\pi}{2} - \theta_n, \frac{\pi}{2}]$, this completes the proof of Lemma 2.1.

A. Proof of Lemma 4.1

We now present the proof of Lemma 4.1.

Proof of Lemma 4.1: Recalling that $\mathbf{Z} = [Z^n(1), Z^n(2), \dots, Z^n(B)]$, we have

$$\|\mathbf{Z}\|^2 = \sum_{b=1}^B \|Z^n(b)\|^2.$$

Therefore by the weak law of large numbers, for any $\delta > 0$ and B sufficiently large we have

$$\Pr \left(\left| \frac{1}{B} \|\mathbf{Z}\|^2 - E[\|Z^n\|^2] \right| \leq \delta \right) \geq 1 - \delta. \quad (45)$$

Since $Z^n = X^n + W_1^n$, we have

$$\begin{aligned} E[\|Z^n\|^2] &= E[\|X^n + W_1^n\|^2] \\ &= E[\|X^n\|^2] + 2E[X^n \cdot W_1^n] + E[\|W_1^n\|^2]. \end{aligned}$$

By assumption $E[\|X^n\|^2] = nP$ and $E[\|W_1^n\|^2] = nN$. Also we have

$$\begin{aligned} E[X^n \cdot W_1^n] &= E\left[\sum_{i=1}^n X_i W_{1i}\right] \\ &= \sum_{i=1}^n E[X_i W_{1i}] \\ &= \sum_{i=1}^n E[X_i]E[W_{1i}] \\ &= 0. \end{aligned}$$

Therefore, $E[\|Z^n\|^2] = n(P + N)$, and this combined with (45) yields that

$$\Pr(\|\mathbf{Z}\|^2 \in [nB(P + N - \delta), nB(P + N + \delta)]) \geq 1 - \delta,$$

for any $\delta > 0$ and B sufficiently large. Since \mathbf{Z} and \mathbf{Y} are identically distributed, the above relation also holds with $\|\mathbf{Z}\|^2$ replaced by $\|\mathbf{Y}\|^2$. This completes the proof of the lemma. \blacksquare

V. PROOF OF LEMMA 4.2

To prove Lemma 4.2, we introduce two lemmas as follows. The proofs of these two lemmas, which will be presented in Sections V-A and V-B, are based on typicality argument and strengthening the isoperimetric inequality on a spherical shell respectively.

Lemma 5.1: Given any pair of (\mathbf{x}, \mathbf{i}) , let $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ be a set of \mathbf{z} defined as

$$\begin{aligned} S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) &:= \left\{ \mathbf{z} \in f^{-1}(\mathbf{i}) : \|\mathbf{x} - \mathbf{z}\| \in [\sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon)] \right. \\ &\quad \left. \mathbf{z} \in \text{Ball}\left(\mathbf{0}, \sqrt{nB(P + N + \epsilon)}\right) \right. \\ &\quad \left. 2^{nB(\log \sin \theta_n - \epsilon)} \leq p(f(\mathbf{z})|\mathbf{x}) \leq 2^{nB(\log \sin \theta_n + \epsilon)} \right\} \end{aligned}$$

where $\theta_n = \arcsin 2^{-\frac{1}{n}H(I_n|X^n)}$. Then there exists a set $S_\epsilon(X^n, I_n)$ of (\mathbf{x}, \mathbf{i}) pairs, such that

$$\Pr((\mathbf{X}, \mathbf{I}) \in S_\epsilon(X^n, I_n)) \geq 1 - \sqrt{\epsilon}, \quad (46)$$

and for any $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$,

$$\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) \geq 2^{nB(\log \sin \theta_n - 2\epsilon)}, \quad (47)$$

for B sufficiently large.

Lemma 5.2: Let $\mathbb{L}^m \subseteq \mathbb{R}^m$ be a spherical shell defined as

$$\mathbb{L}^m = \{\mathbf{y} \in \mathbb{R}^m : R_m - \Delta \leq \|\mathbf{y}\| \leq R_m + \Delta\}$$

where $R_m := \sqrt{mR}$ for some $R > 0$ and $\Delta := \sqrt{m}\delta$ for some $0 < \delta \leq \sqrt{R}$. Let $A \subseteq \mathbb{L}^m$ be an arbitrary subset on this shell with volume

$$|A| = 2^{\frac{m}{2} \log 2\pi e R \sin^2 \theta}.$$

For any $\omega \in (\pi/2 - \theta, \pi/2]$, $\epsilon > 0$ and m sufficiently large, we have

$$\Pr\left(\left|A \cap \text{Ball}\left(\mathbf{Y}, 2(R_m + \Delta)\sin \frac{\omega}{2} + 2\Delta\right)\right| \geq 2^{m[\frac{1}{2} \log(2\pi e R(\sin^2 \theta - \cos^2 \omega)) - \epsilon]}\right) \geq 1 - \epsilon, \quad (48)$$

where \mathbf{Y} is uniformly drawn from \mathbb{L}^m .

Equipped with the above two lemmas, we are now ready to prove Lemma 4.2.

Proof of Lemma 4.2: From Lemma 5.1, we have for any pair $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$,

$$\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) \geq 2^{nB(\log \sin \theta_n - 2\epsilon)},$$

for B sufficiently large. We also have

$$\begin{aligned} \Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) &\leq |S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})| \sup_{\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})} f(\mathbf{z}|\mathbf{x}) \\ &\leq |S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})| 2^{-nB(\frac{1}{2} \log 2\pi e N - \epsilon_1)}, \end{aligned}$$

for some $\epsilon_1 \rightarrow 0$ as $\epsilon \rightarrow 0$, where the second inequality follows because i) \mathbf{Z} is Gaussian distributed given \mathbf{x} , and ii) for any $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$, we have

$$\|\mathbf{x} - \mathbf{z}\| \in [\sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon)].$$

Therefore, the volume of $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ can be lower bounded by

$$|S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})| \geq 2^{nB(\frac{1}{2} \log(2\pi e N \sin^2 \theta_n) - 2\epsilon - \epsilon_1)}.$$

Noting that $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ is a subset of

$$\text{Shell}(\mathbf{x}, \sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon)),$$

by Lemma 5.2, we have for any $\omega \in (\pi/2 - \theta_n, \pi/2]$,

$$\Pr\left(\left|S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) \cap \text{Ball}\left(\mathbf{U}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \epsilon_2\right)}\right)\right| \geq 2^{nB\left[\frac{1}{2} \log(2\pi e N(\sin^2 \theta_n - \cos^2 \omega)) - \epsilon_3\right]}\right) \geq 1 - \epsilon_3 \quad (49)$$

where \mathbf{U} is uniformly distributed over $\text{Shell}(\mathbf{x}, \sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon))$, ϵ_2 is defined such that

$$\sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \epsilon_2\right)} = 2\left(\sqrt{nBN} + \sqrt{nB\epsilon}\right) \sin \frac{\omega}{2} + 2\sqrt{nB\epsilon},$$

which tends to zero as ϵ goes to zero, and $\epsilon_3 \rightarrow 0$ as $B \rightarrow \infty$.

Now we translate the bound (49) on the probability involving \mathbf{U} to a bound on the probability involving \mathbf{Y} . Specifically, let $\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}$ be defined as

$$\begin{aligned} \mathcal{Y}_{(\mathbf{x}, \mathbf{i})} = &\left\{ \mathbf{y} \in \text{Shell}(\mathbf{x}, \sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon)) : \right. \\ &\left. \left|S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) \cap \text{Ball}\left(\mathbf{U}, \sqrt{nBN\left(4\sin^2\frac{\omega}{2} + \epsilon_2\right)}\right)\right| \geq 2^{nB\left[\frac{1}{2} \log(2\pi e N(\sin^2 \theta_n - \cos^2 \omega)) - \epsilon_3\right]} \right\}. \end{aligned}$$

By (49), the volume of $\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}$ is lower bounded by

$$\begin{aligned} |\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}| &\geq (1 - \epsilon_3) \left| \text{Shell}(\mathbf{x}, \sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon)) \right| \\ &\geq 2^{nB(\frac{1}{2} \log 2\pi e N - \epsilon_4)}, \end{aligned}$$

where $\epsilon_4 \rightarrow 0$ as $\epsilon \rightarrow 0$ and $B \rightarrow \infty$. Therefore, we have

$$\begin{aligned} \Pr(\mathbf{Y} \in \mathcal{Y}_{(\mathbf{x}, \mathbf{i})}|\mathbf{x}) &\geq |\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}| \inf_{\mathbf{y} \in \mathcal{Y}_{(\mathbf{x}, \mathbf{i})}} f(\mathbf{y}|\mathbf{x}) \\ &\geq 2^{nB(\frac{1}{2} \log 2\pi e N - \epsilon_4)} 2^{-nB(\frac{1}{2} \log 2\pi e N + \epsilon_1)} \\ &\geq 2^{-nB\epsilon_5}, \end{aligned}$$

where $\epsilon_5 \rightarrow 0$ as $\epsilon \rightarrow 0$ and $B \rightarrow \infty$. We then apply the following blowing-up lemma to the set $\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}$; for a proof of this lemma, see [6].

Lemma 5.3: Let V_1, V_2, \dots, V_n be n i.i.d. Gaussian random variables with $V_i \sim \mathcal{N}(0, N)$, $\forall i \in \{1, 2, \dots, n\}$. Then, for any $A \subseteq \mathbb{R}^n$ with $\Pr(V^n \in A) \geq 2^{-na_n}$,

$$\Pr(V^n \in \Gamma_{\sqrt{n}(\sqrt{2Na_n \ln 2} + r)}(A)) \geq 1 - 2^{-\frac{nr^2}{2N}}, \forall r > 0,$$

where

$$\Gamma_l(A) := \{v^n \in \mathbb{R}^n : \exists v_1^n \in A \text{ s.t. } \|v^n - v_1^n\| \leq l\}.$$

In particular, noting that \mathbf{Y} is Gaussian distributed given \mathbf{x} , the above lemma immediately yields that for $\epsilon_6 := \sqrt{2N\epsilon_5 \ln 2} + \epsilon$,

$$\begin{aligned} \Pr(\mathbf{Y} \in \Gamma_{\sqrt{nB}\epsilon_6}(\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}) | \mathbf{x}) &\geq 1 - 2^{-\frac{nB\epsilon^2}{2N}} \\ &\geq 1 - \epsilon, \end{aligned}$$

for B sufficiently large.

Let ϵ_7 be defined such that

$$\sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_7\right)} = \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_2\right)} + \sqrt{nB}\epsilon_6.$$

Then we have for any $\mathbf{y} \in \Gamma_{\sqrt{nB}\epsilon_6}(\mathcal{Y}_{(\mathbf{x}, \mathbf{i})})$, there exists a $\mathbf{y}' \in \mathcal{Y}_{(\mathbf{x}, \mathbf{i})}$ such that $\|\mathbf{y} - \mathbf{y}'\| \leq \sqrt{nB}\epsilon_6$ and

$$\text{Ball} \left(\mathbf{y}', \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_2\right)} \right) \subseteq \text{Ball} \left(\mathbf{y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_7\right)} \right).$$

This implies that for any $\mathbf{y} \in \Gamma_{\sqrt{nB}\epsilon_6}(\mathcal{Y}_{(\mathbf{x}, \mathbf{i})})$,

$$\left| S_\epsilon(Z^n | \mathbf{x}, \mathbf{i}) \cap \text{Ball} \left(\mathbf{y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_7\right)} \right) \right| \geq 2^{nB \left[\frac{1}{2} \log(2\pi e N (\sin^2 \theta_n - \cos^2 \omega)) - \epsilon_3 \right]}.$$

By the definition $S_\epsilon(Z^n | \mathbf{x}, \mathbf{i})$ is a subset of $f^{-1}(\mathbf{i}) \cap \text{Ball} \left(\mathbf{0}, \sqrt{nB(P + N + \epsilon)} \right)$, therefore we in turn have

$$\left| f^{-1}(\mathbf{i}) \cap \text{Ball} \left(\mathbf{0}, \sqrt{nB(P + N + \epsilon)} \right) \cap \text{Ball} \left(\mathbf{y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_7\right)} \right) \right| \geq 2^{nB \left[\frac{1}{2} \log(2\pi e N (\sin^2 \theta_n - \cos^2 \omega)) - \epsilon_3 \right]}.$$

for any $\mathbf{y} \in \Gamma_{\sqrt{nB}\epsilon_6}(\mathcal{Y}_{(\mathbf{x}, \mathbf{i})})$.

Therefore, we have for B sufficiently large,

$$\begin{aligned} &\Pr \left(\left| f^{-1}(\mathbf{I}) \cap \text{Ball} \left(\mathbf{0}, \sqrt{nB(P + N + \epsilon)} \right) \cap \text{Ball} \left(\mathbf{Y}, \sqrt{nBN \left(4\sin^2 \frac{\omega}{2} + \epsilon_7\right)} \right) \right| \right. \\ &\quad \left. \geq 2^{nB \left[\frac{1}{2} \log(2\pi e N (\sin^2 \theta_n - \cos^2 \omega)) - \epsilon_3 \right]} \right) \\ &\geq \sum_{(\mathbf{x}, \mathbf{i})} \Pr(\mathbf{Y} \in \Gamma_{\sqrt{nB}\epsilon_6}(\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}) | \mathbf{x}) p(\mathbf{x}, \mathbf{i}) \\ &\geq \sum_{(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)} \Pr(\mathbf{Y} \in \Gamma_{\sqrt{nB}\epsilon_6}(\mathcal{Y}_{(\mathbf{x}, \mathbf{i})}) | \mathbf{x}) p(\mathbf{x}, \mathbf{i}) \\ &\geq (1 - \epsilon)(1 - \sqrt{\epsilon}) \\ &\geq 1 - 2\sqrt{\epsilon}. \end{aligned}$$

Finally, choosing $\delta = \max\{2\sqrt{\epsilon}, \epsilon_3, \epsilon_7\}$ concludes the proof of Lemma 4.2. ■

A. Proof of Lemma 5.1 Via Typicality Argument

We now present the proof of Lemma 5.1. Recall that $H(I_n|X^n) = -n \log \sin \theta_n$ for the n -channel use code. By the law of large numbers and Lemma 4.1, we have for any $\epsilon > 0$ and sufficiently large B ,

$$\Pr((\mathbf{X}, \mathbf{Z}) \in S_\epsilon(X^n, Z^n)) \geq 1 - \epsilon$$

where

$$\begin{aligned} S_\epsilon(X^n, Z^n) := & \left\{ (\mathbf{x}, \mathbf{z}) : \|\mathbf{x} - \mathbf{z}\| \in [\sqrt{nB}(\sqrt{N} - \epsilon), \sqrt{nB}(\sqrt{N} + \epsilon)] \right. \\ & \mathbf{z} \in \text{Ball}\left(\mathbf{0}, \sqrt{nB(P + N + \epsilon)}\right) \\ & \left. 2^{nB(\log \sin \theta_n - \epsilon)} \leq p(f(\mathbf{z})|\mathbf{x}) \leq 2^{nB(\log \sin \theta_n + \epsilon)} \right\}. \end{aligned}$$

Note that in terms of $S_\epsilon(X^n, Z^n)$, the set $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ in Lemma 5.1 can be simply written as

$$S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}) = \{\mathbf{z} : f(\mathbf{z}) = \mathbf{i}, (\mathbf{x}, \mathbf{z}) \in S_\epsilon(X^n, Z^n)\}.$$

Therefore, for B sufficiently large, we have

$$\begin{aligned} \Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{X}, \mathbf{I})) &= \Pr(f(\mathbf{Z}) = \mathbf{I}, (\mathbf{X}, \mathbf{Z}) \notin S_\epsilon(X^n, Z^n)) \\ &\leq \epsilon. \end{aligned}$$

On the other hand, defining $S_\epsilon(X^n, I_n) := \{(\mathbf{x}, \mathbf{i}) : \Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) \geq 1 - \sqrt{\epsilon}\}$, we have

$$\begin{aligned} \Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{X}, \mathbf{I})) &= \sum_{(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)} \Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i})p(\mathbf{x}, \mathbf{i}) \\ &\quad + \sum_{(\mathbf{x}, \mathbf{i}) \notin S_\epsilon(X^n, I_n)} \Pr(\mathbf{Z} \notin S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i})p(\mathbf{x}, \mathbf{i}) \\ &\geq \sqrt{\epsilon} \cdot \Pr(S_\epsilon^c(X^n, I_n)). \end{aligned}$$

Therefore, we have for B sufficiently large,

$$\Pr(S_\epsilon^c(X^n, I_n)) \leq \frac{\epsilon}{\sqrt{\epsilon}} = \sqrt{\epsilon},$$

and thus

$$\Pr(S_\epsilon(X^n, I_n)) \geq 1 - \sqrt{\epsilon},$$

which proves (46).

To prove (47), consider any $(\mathbf{x}, \mathbf{i}) \in S_\epsilon(X^n, I_n)$. From the definition of $S_\epsilon(X^n, I_n)$, $\Pr(S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) \geq 1 - \sqrt{\epsilon}$. Therefore, $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$ must be nonempty, i.e., there exists at least one $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$. Consider any $\mathbf{z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$. By the definition of $S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})$, we have $f(\mathbf{z}) = \mathbf{i}$ and $(\mathbf{x}, \mathbf{z}) \in S_\epsilon(X^n, Z^n)$. Then, it follows from the definition of $S_\epsilon(X^n, Z^n)$ that

$$2^{nB(\log \sin \theta_n - \epsilon)} \leq p(f(\mathbf{z})|\mathbf{x}) = p(\mathbf{i}|\mathbf{x}) \leq 2^{nB(\log \sin \theta_n + \epsilon)}.$$

This further implies that

$$\begin{aligned} & \Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}) \\ &= \frac{\Pr(f(\mathbf{Z}) = \mathbf{i}|\mathbf{x})\Pr(\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, f(\mathbf{Z}) = \mathbf{i})}{\Pr(f(\mathbf{Z}) = \mathbf{i}|\mathbf{Z} \in S_\epsilon(Z^n|\mathbf{x}, \mathbf{i}), \mathbf{x})} \\ &= p(\mathbf{i}|\mathbf{x})\Pr(S_\epsilon(Z^n|\mathbf{x}, \mathbf{i})|\mathbf{x}, \mathbf{i}) \\ &\geq 2^{nB(\log \sin \theta_n - \epsilon)}(1 - \sqrt{\epsilon}) \\ &\geq 2^{nB(\log \sin \theta_n - 2\epsilon)} \end{aligned}$$

for sufficiently large B , which concludes the proof of (47) and Lemma 5.1.

B. Proof of Lemma 5.2 Via Strengthening Isoperimetric Inequality

We now prove Lemma 5.2 by extending our result in Lemma 3.1 on a sphere to a spherical shell. Let

$$\mathbb{L}^m = \{\mathbf{y} \in \mathbb{R}^m : R_m - \Delta \leq \|\mathbf{y}\| \leq R_m + \Delta\}$$

be this shell where $R_m := \sqrt{mR}$ for some $R > 0$ and $\Delta := \sqrt{m}\delta$ for some $0 < \delta \leq \sqrt{R}$; note that one can also write this shell as

$$\mathbb{L}^m = \bigcup_{r \in [R_m - \Delta, R_m + \Delta]} \mathbb{S}_r^{m-1}$$

where

$$\mathbb{S}_r^{m-1} := \{\mathbf{y} \in \mathbb{R}^m : \|\mathbf{y}\| = r\}$$

denotes the $(m-1)$ -sphere of radius r .

Let $A \subseteq \mathbb{L}^m$ be an arbitrary subset on this shell with volume

$$|A| = 2^{\frac{m}{2} \log 2\pi e R \sin^2 \theta}.$$

We say A is spherically symmetric about a pole $\mathbf{z}_0 \in \mathbb{S}_{R_m}^{m-1}$ if for any $r \in [R_m - \Delta, R_m + \Delta]$, the intersection $A \cap \mathbb{S}_r^{m-1}$ yields a spherical cap about the pole $r\mathbf{z}_0/R_m$. In this case, denote by θ_r the angle associated with each cap $A \cap \mathbb{S}_r^{m-1}$.

Similarly as the proof of Lemma 3.1, we prove Lemma 5.2 in two steps. In particular, we will first demonstrate that the claim in Lemma 5.2 holds when A is known to be spherically symmetric, and then show that if it holds for a spherically symmetric A , then it holds for any arbitrary set A .

i) Spherically symmetric A : Suppose A is spherically symmetric about pole \mathbf{z}_0 . With a slight abuse of notation, denote by

$$\text{Cap}(\mathbf{y}, \omega) := \{\mathbf{y}' \in \mathbb{L}^m : \angle(\mathbf{y}, \mathbf{y}') \leq \omega\}$$

a shell cap in \mathbb{L}^m with pole \mathbf{y} and angle ω . To show (48), it suffices to show that for any $\omega > \pi/2 - \theta$, $\varepsilon > 0$ and m sufficiently large,

$$\Pr \left(|A \cap \text{Cap}(\mathbf{Y}, \omega)| \geq 2^{m \left(\frac{1}{2} \log 2\pi e R (\sin^2 \theta - \cos^2 \omega) - \varepsilon \right)} \right) > 1 - \varepsilon$$

where \mathbf{Y} is uniformly distributed over \mathbb{L}^m . This is because the shell $\text{Cap}(\mathbf{Y}, \omega)$ is always contained by the Euclidean Ball $(\mathbf{Y}, 2(R_m + \Delta) \sin \frac{\omega}{2} + 2\Delta)$ and thus

$$A \cap \text{Cap}(\mathbf{Y}, \omega) \subseteq A \cap \text{Ball} \left(\mathbf{Y}, 2(R_m + \Delta) \sin \frac{\omega}{2} + 2\Delta \right).$$

Since the intersection volume $|A \cap \text{Cap}(\mathbf{y}, \omega)|$ depends only on the angle $\angle(\mathbf{y}, \mathbf{y}_0)$ and not on the magnitude of \mathbf{y} , we can instead consider \mathbf{y} to be uniformly drawn from the sphere $\mathbb{S}_{R_m}^{m-1}$. For each $\mathbf{y} \in \mathbb{S}_{R_m}^{m-1}$, the intersection volume $|A \cap \text{Cap}(\mathbf{y}, \omega)|$ can be written as an integral over each sphere in the shell, i.e.,

$$\begin{aligned} |A \cap \text{Cap}(\mathbf{y}, \omega)| &= \int_{\mathbb{R}^m} 1_{A \cap \text{Cap}(\mathbf{y}, \omega)}(\mathbf{z}) d\mathbf{z} \\ &= \int_{R_m - \Delta}^{R_m + \Delta} \left(\int_{\mathbb{S}_r^{m-1}} 1_{A \cap \text{Cap}(\mathbf{y}, \omega)}(\mathbf{z}) d\mathbf{z} \right) dr. \end{aligned}$$

Recall that due to measure concentration, for any $\varepsilon > 0$ and m sufficiently large, we have

$$\Pr (\angle(\mathbf{z}_0, \mathbf{Y}) \in [\pi/2 - \varepsilon/2, \pi/2 + \varepsilon/2]) \geq 1 - \varepsilon. \quad (50)$$

This allows us to concentrate on characterizing $|A \cap \text{Cap}(\mathbf{y}, \omega)|$ for \mathbf{y} 's in $\mathbb{S}_{R_m}^{m-1}$ with $\angle(\mathbf{z}_0, \mathbf{y}) \in [\pi/2 - \varepsilon/2, \pi/2 + \varepsilon/2]$.

In particular, for $\mathbf{y}_0 \in \mathbb{S}_{R_m}^{m-1}$ with $\angle(\mathbf{z}_0, \mathbf{y}_0) = \pi/2$ and any $r \in [R_m - \Delta, R_m + \Delta]$, it can be shown that

$$\int_{\mathbb{S}_r^{m-1}} 1_{A \cap \text{Cap}(\mathbf{y}_0, \omega)}(\mathbf{z}) d\mathbf{z} \quad (51)$$

$$\geq 2^m \left(\frac{1}{2} \log 2\pi e R'_r (\sin^2 \theta_r - \cos^2 \omega) - \epsilon \right) \quad (52)$$

$$\geq 2^m \left(\frac{1}{2} \log 2\pi e R' (\sin^2 \theta_r - \cos^2 \omega) - \epsilon \right) \quad (53)$$

where $\omega > \pi/2 - \theta_r$ and ϵ can be made arbitrarily small by choosing m to be sufficiently large. In the above, (52) follows from characterizing the surface area of the intersection of two spherical caps (see Appendix B), where R'_r is defined such that $\sqrt{m R'_r} = r$, and (53) follows from lower bounding R'_r by R' with R' being defined such that $\sqrt{m R'} = R_m - \Delta$. We would like to integrate (53) over r in order to get a lower bound on the total intersection volume $|A \cap \text{Cap}(\mathbf{y}_0, \omega)|$. However, (53) is only valid when $\theta_r + \omega \geq \frac{\pi}{2}$, and when $\theta_r + \omega < \frac{\pi}{2}$, the slice intersection area (51) will be empty and does not contribute to the total intersection volume. This motivates us to set

$$\theta'_r = \max \left(\theta_r, \frac{\pi}{2} - \omega \right)$$

and write

$$\int_{\mathbb{S}_r^{m-1}} 1_{A \cap \text{Cap}(\mathbf{y}_0, \omega)}(\mathbf{z}) d\mathbf{z} \geq 2^m \left(\frac{1}{2} \log 2\pi e R' (\sin^2 \theta'_r - \cos^2 \omega) - \epsilon \right). \quad (54)$$

We now rewrite the right-hand side of (54) in the following way. Let the function $h : [0, 1] \rightarrow \mathbb{R}$ be defined by

$$h(t) = 2^{\frac{m}{2} \log 2\pi e R' (t^{\frac{2}{m}} - \cos^2 \omega)}$$

so that

$$h(\sin^m \theta'_r) = 2^{\frac{m}{2} \log 2\pi e R' (\sin^2 \theta'_r - \cos^2 \omega)}.$$

By differentiating $h(t)$ twice it can be seen that h is convex for $t^{\frac{2}{m}} \geq \cos^2 \omega$. Integrating both sides of (54) yields that for any $\epsilon > 0$ and m sufficiently large,

$$\begin{aligned} & \int_{R_m - \Delta}^{R_m + \Delta} \left(\int_{\mathbb{S}_r^{m-1}} 1_{A \cap \text{Cap}(\mathbf{y}_0, \omega)}(\mathbf{z}) d\mathbf{z} \right) dr \\ & \geq \int_{R_m - \Delta}^{R_m + \Delta} h(\sin^m \theta'_r) 2^{-m\epsilon} dr \end{aligned} \quad (55)$$

$$\geq 2\Delta h \left(\frac{1}{2\Delta} \int_{R_m - \Delta}^{R_m + \Delta} \sin^m \theta'_r dr \right) 2^{-m\epsilon} \quad (56)$$

$$\geq 2^{-m\epsilon} h \left(\frac{1}{2\Delta} \int_{R_m - \Delta}^{R_m + \Delta} \sin^m \theta'_r dr \right) \quad (57)$$

where (56) is due to Jensen's inequality.

On the other hand, we have

$$\begin{aligned} |A| &= 2^{m \left(\frac{1}{2} \log 2\pi e R \sin^2 \theta \right)} \\ &\leq \int_{R_m - \Delta}^{R_m + \Delta} 2^{m \left(\frac{1}{2} \log 2\pi e R'_r \sin^2 \theta_r + \epsilon \right)} dr \\ &\leq \int_{R_m - \Delta}^{R_m + \Delta} 2^{m \left(\frac{1}{2} \log 2\pi e R'' \sin^2 \theta_r + \epsilon \right)} dr \\ &= 2^{m \left(\frac{1}{2} \log 2\pi e R'' + \epsilon \right)} \int_{R_m - \Delta}^{R_m + \Delta} \sin^m \theta_r dr \end{aligned}$$

where the first inequality follows from characterizing the surface area of a hyperspherical cap (see Appendix A), and the second inequality follows by defining R'' to be such that $\sqrt{mR''} = R_m + \Delta$. Therefore, we have

$$\begin{aligned} \frac{1}{2\Delta} \int_{R_m-\Delta}^{R_m+\Delta} \sin^m \theta_r dr &\geq \frac{1}{2\Delta} 2^m \left(\frac{1}{2} \log \left(\frac{R \sin^2 \theta}{R''} \right) - \epsilon \right) \\ &\geq 2^m \left(\frac{1}{2} \log \sin^2 \theta - 2\epsilon \right), \end{aligned}$$

for any $\epsilon > 0$ and m sufficiently large. Combining this with (57), we have for m sufficiently large,

$$\begin{aligned} |A \cap \text{Cap}(\mathbf{y}_0, \omega)| &\geq 2^{-m\epsilon} h \left(\frac{1}{2\Delta} \int_{R_m-\Delta}^{R_m+\Delta} \sin^m \theta'_r dr \right) \\ &\geq 2^{-m\epsilon} h \left(\frac{1}{2\Delta} \int_{R_m-\Delta}^{R_m+\Delta} \sin^m \theta_r dr \right) \\ &\geq 2^{-m\epsilon} h \left(2^m \left(\frac{1}{2} \log \sin^2 \theta - 2\epsilon \right) \right) \\ &= 2^{-m\epsilon} 2^{\frac{m}{2} \log 2\pi e R' (\sin^2 \theta - 2^{-4\epsilon} - \cos^2 \omega)} \\ &\geq 2^{m \left(\frac{1}{2} \log 2\pi e R (\sin^2 \theta - \cos^2 \omega) - \epsilon_1 \right)} \end{aligned}$$

where $\epsilon, \epsilon_1 \rightarrow 0$ as $m \rightarrow \infty$. This together with (50) then yields that

$$\Pr \left(|A \cap \text{Cap}(\mathbf{Y}, \omega)| \geq 2^{m \left(\frac{1}{2} \log 2\pi e R (\sin^2 \theta - \cos^2 \omega) - \epsilon \right)} \right) \geq 1 - \epsilon, \quad (58)$$

where $\epsilon \rightarrow 0$ as $m \rightarrow \infty$.

ii) Arbitrary A : To extend the above result to an arbitrary A , we again use the Riesz rearrangement inequality in Lemma 3.2, similarly as in the proof of Lemma 3.1. In particular, to apply Lemma 3.2 to the shell situation we must find appropriate functions on the sphere. Note that

$$\begin{aligned} |A \cap \text{Cap}(\mathbf{y}, \omega)| &= \int_{\mathbb{R}^m} 1_{A \cap \text{Cap}(\mathbf{y}, \omega)}(\mathbf{z}) d\mathbf{z} \\ &= \int_{\mathbb{S}_1^{m-1}} \left(\int_{R_m-\Delta}^{R_m+\Delta} r^{m-1} 1_{A \cap \text{Cap}(\mathbf{y}, \omega)}(r\mathbf{z}) dr \right) d\mathbf{z} \end{aligned}$$

so that if

$$f_A(\mathbf{z}) = \int_{R_m-\Delta}^{R_m+\Delta} r^{m-1} 1_A(r\mathbf{z}) dr$$

and

$$K(\cos \alpha) = \begin{cases} 0 & \omega < \alpha \leq \pi \\ 1 & 0 \leq \alpha \leq \omega \end{cases}$$

then

$$\psi(\mathbf{y}) = |A \cap \text{Cap}(\mathbf{y}, \omega)| = \int_{\mathbb{S}_1^{m-1}} f_A(\mathbf{z}) K(\langle \mathbf{z}, \mathbf{y} / \|\mathbf{y}\| \rangle) d\mathbf{z}.$$

Let ψ^* be the symmetric decreasing rearrangement of ψ about a pole \mathbf{z}_0 . The claim is that there exists a subset $A' \subseteq \mathbb{L}^m$ with $|A| = |A'|$ which is spherically symmetric about \mathbf{z}_0 and such that $f_{A'} = f_A^*$. Explicitly, one example would be

$$A' = \bigcup_{r \in [R_m-\Delta, R_m+\Delta]} A'_r$$

where $A'_r \subseteq \mathbb{S}_r^{m-1}$ is a spherical cap with the same area as $\{\mathbf{z} : f_A(\mathbf{z}) > d\}$ for $d = \int_{R_m-\Delta}^r r^{m-1} dr$.

Let $\psi'(\mathbf{y}) = |A' \cap \text{Cap}(\mathbf{y}, \omega)|$. Both ψ and ψ' depend only on the angle $\angle(\mathbf{z}_0, \mathbf{y})$ and are nonincreasing in that angle. Following the same argument from the proof of Lemma 3.1, one can establish (58) for any arbitrary $A \subseteq \mathbb{L}^m$, which finishes the proof of Lemma 5.2.

APPENDIX A

SURFACE AREA AND VOLUME OF A HYPERSPHERICAL CAP

In this Appendix we derive the surface area and volume formulas for a hyperspherical cap (see also [13]) and then characterize the exponents of these expressions.

Let $C \subseteq \mathbb{S}^{m-1}$ be a spherical cap with angle θ on the $(m-1)$ -sphere of radius $R_m = \sqrt{mR}$. The area $\mu(C)$ of C can be written as

$$\mu(C) = \int_0^\theta A_{m-2}(R_m \sin \rho) R_m d\rho$$

where $A_{m-2}(R_m \sin \rho)$ is the total surface area of the $(m-2)$ -sphere of radius $R_m \sin \rho$. Plugging in the expression for the surface area of an $(m-2)$ -sphere leads to

$$\mu(C) = \frac{2\pi^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})} (mR)^{\frac{m-2}{2}} \int_0^\theta \sin^{m-2} \rho d\rho.$$

We now characterize the exponent of $\mu(C)$. First, by Stirling's approximation, $\frac{2\pi^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})} (mR)^{\frac{m-2}{2}}$ in the above can be bounded as

$$2^{\frac{m}{2} [\log(2\pi e R) - \epsilon_1]} \leq \frac{2\pi^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})} (mR)^{\frac{m-2}{2}} \leq 2^{\frac{m}{2} [\log(2\pi e R) + \epsilon_1]}$$

for some $\epsilon_1 \rightarrow 0$ as $m \rightarrow \infty$. Also for m sufficiently large, we have

$$\begin{aligned} \int_0^\theta \sin^{m-2} \rho d\rho &= \int_0^\theta 2^{\frac{m-2}{2} \log \sin^2 \rho} d\rho \\ &\geq \int_{\theta - \frac{1}{m}}^\theta 2^{\frac{m-2}{2} \log \sin^2 \rho} d\rho \\ &\geq \frac{1}{m} 2^{\frac{m-2}{2} \log \sin^2(\theta - \frac{1}{m})} \\ &\geq 2^{\frac{m}{2} (\log \sin^2 \theta - \epsilon_2)} \end{aligned}$$

and

$$\begin{aligned} \int_0^\theta \sin^{m-2} \rho d\rho &= \int_0^\theta 2^{\frac{m-2}{2} \log \sin^2 \rho} d\rho \\ &\leq \theta \cdot 2^{\frac{m-2}{2} \log \sin^2 \theta} \\ &\leq 2^{\frac{m}{2} (\log \sin^2 \theta + \epsilon_2)} \end{aligned}$$

for some $\epsilon_2 \rightarrow 0$ as $m \rightarrow \infty$. Therefore, the area $\mu(C)$ can be bounded as

$$2^{\frac{m}{2} [\log(2\pi e R \sin^2 \theta) - \epsilon]} \leq \mu(C) \leq 2^{\frac{m}{2} [\log(2\pi e R \sin^2 \theta) + \epsilon]} \quad (59)$$

for some $\epsilon \rightarrow 0$ as $m \rightarrow \infty$.

To find the volume $|C|$ of C , we integrate over the surface area of the spheres at each radius, which leads to

$$\begin{aligned} |C| &= \frac{2\pi^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})} \int_0^\theta \sin^{m-2} \rho d\rho \int_0^{R_m} r^{m-2} dr \\ &= \frac{2(\pi m R)^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2}) (m-1)} \int_0^\theta \sin^{m-2} \rho d\rho. \end{aligned}$$

Following the similar approach as above, we can also bound the volume $|C|$ as

$$2^{\frac{m}{2} [\log(2\pi e R \sin^2 \theta) - \epsilon]} \leq |C| \leq 2^{\frac{m}{2} [\log(2\pi e R \sin^2 \theta) + \epsilon]} \quad (60)$$

for some $\epsilon \rightarrow 0$ as $m \rightarrow \infty$.

APPENDIX B

SURFACE AREA OF THE INTERSECTION OF TWO HYPERSPHERICAL CAPS

Let $\mathbb{S}^{m-1} \subseteq \mathbb{R}^m$ be an $(m-1)$ -sphere of radius $R_m = \sqrt{mR}$. Let

$$C_i = \text{Cap}(\mathbf{v}_i, \theta_i) = \{\mathbf{v} \in \mathbb{S}^{m-1} | \angle(\mathbf{v}, \mathbf{v}_i) \leq \theta_i\}, i = 1, 2$$

be two spherical caps on \mathbb{S}^{m-1} such that $\angle(\mathbf{v}_1, \mathbf{v}_2) = \frac{\pi}{2}$, $\theta_i < \frac{\pi}{2}$, and $\theta_1 + \theta_2 > \frac{\pi}{2}$. We have the following lemma which characterizes the intersection area $\mu(C_1 \cap C_2)$ of these two caps.

Lemma B.1: There exists some ϵ which tends to 0 as m goes to infinity, such that

$$\begin{aligned} \mu(C_1 \cap C_2) &\geq 2^{\frac{m}{2} [\log(2\pi e R (\sin^2 \theta_1 - \cos^2 \theta_2)) - \epsilon]} \\ \text{and } \mu(C_1 \cap C_2) &\leq 2^{\frac{m}{2} [\log(2\pi e R (\sin^2 \theta_1 - \cos^2 \theta_2)) + \epsilon]}. \end{aligned}$$

To prove this lemma, below we will first derive the surface area formula for the intersection of the above two caps (see also [14]) and then characterize the exponent of this area.

A. Deriving the Surface Area Formula

Consider the points $\mathbf{v} \in \mathbb{S}^{m-1}$ such that

$$\angle(\mathbf{v}_1, \mathbf{v}) = \theta_1$$

and

$$\angle(\mathbf{v}_2, \mathbf{v}) = \theta_2.$$

These points satisfy the linear relations

$$\langle \mathbf{v}_1, \mathbf{v} \rangle = R_m^2 \cos \theta_1$$

and

$$\langle \mathbf{v}_2, \mathbf{v} \rangle = R_m^2 \cos \theta_2,$$

and therefore all such \mathbf{v} lie in the unique $m-1$ dimensional subspace H defined by

$$\left\langle \frac{\mathbf{v}_1}{\cos \theta_1} - \frac{\mathbf{v}_2}{\cos \theta_2}, \mathbf{v} \right\rangle = 0.$$

The angle between the hyperplane H and the vector \mathbf{v}_2 is

$$\phi = \frac{\pi}{2} - \arccos \left(\frac{1}{R_m \sqrt{\frac{1}{\cos^2 \theta_1} + \frac{1}{\cos^2 \theta_2}}} \left\langle \frac{\mathbf{v}_1}{\cos \theta_1} - \frac{\mathbf{v}_2}{\cos \theta_2}, \mathbf{v}_2 \right\rangle \right)$$

and because \mathbf{v}_1 and \mathbf{v}_2 are orthogonal and $\|\mathbf{v}_2\| = R_m$,

$$\phi = \frac{\pi}{2} - \arccos \left(\frac{1}{\cos \theta_2 \sqrt{\frac{1}{\cos^2 \theta_1} + \frac{1}{\cos^2 \theta_2}}} \right) = \arctan \left(\frac{\cos \theta_1}{\cos \theta_2} \right).$$

The approach will be as follows. Divide the intersection $C_1 \cap C_2$ into two parts C^+ and C^- that are on either side of the hyperplane H . More concretely,

$$C^+ = \left\{ \mathbf{v} \in C_1 \cap C_2 \mid \left\langle \mathbf{v}, \frac{\mathbf{v}_1}{\cos \theta_1} - \frac{\mathbf{v}_2}{\cos \theta_2} \right\rangle \geq 0 \right\}$$

and

$$C^- = \left\{ \mathbf{v} \in C_1 \cap C_2 \mid \left\langle \mathbf{v}, \frac{\mathbf{v}_1}{\cos \theta_1} - \frac{\mathbf{v}_2}{\cos \theta_2} \right\rangle < 0 \right\}.$$

Each part C^+ and C^- can be written as a union of lower dimensional spherical caps. We will find the measure of each part by integrating the measures of these lower dimensional caps.

The measure of the cap C_2 can be expressed as the integral

$$\mu(C_2) = \int_0^{\theta_2} A_{m-2}(R_m \sin \rho) R d\rho$$

where $A_{m-2}(R_m \sin \rho)$ is the surface area of the $(m-2)$ -sphere with radius $R_m \sin \rho$. If we consider a single $(m-2)$ -sphere at some angle ρ , then the hyperplane H divides that $(m-2)$ -sphere into two spherical caps. The claim is that each of these $m-2$ dimensional caps that is on the side of H with \mathbf{v}_1 is contained in C^+ (and those on the side with \mathbf{v}_2 are contained in C^-). Furthermore, all points in C^+ are in one of these $m-2$ dimensional caps. The claim follows because

$$\left\langle \mathbf{v}, \frac{\mathbf{v}_1}{\cos \theta_1} - \frac{\mathbf{v}_2}{\cos \theta_2} \right\rangle \geq 0$$

implies

$$\cos \theta_2 \cos (\angle(\mathbf{v}, \mathbf{v}_1)) \geq \cos \theta_1 \cos (\angle(\mathbf{v}, \mathbf{v}_2))$$

and since $\angle(\mathbf{v}, \mathbf{v}_2) \leq \theta_2$ and $\cos (\angle(\mathbf{v}, \mathbf{v}_2)) \geq \cos \theta_2$, this implies

$$\cos \theta_2 \cos (\angle(\mathbf{v}, \mathbf{v}_1)) \geq \cos \theta_1 \cos \theta_2.$$

Finally, this implies $\angle(\mathbf{v}, \mathbf{v}_1) \leq \theta_1$, $\mathbf{v} \in C_1$, and $\mathbf{v} \in C^+$.

Note that for $\rho < \phi$, the $(m-2)$ -sphere at angle ρ is entirely on the \mathbf{v}_2 side of H . This establishes the fact that

$$\mu(C^+) = \int_{\phi}^{\theta_2} C_{m-2}^{\theta_{\rho}}(R_m \sin \rho) R d\rho$$

where $C_{m-2}^{\theta_{\rho}}(R_m \sin \rho)$ is the surface area of an $m-2$ dimensional spherical cap defined by angle θ_{ρ} on the $(m-2)$ -sphere of radius $R_m \sin \rho$. Writing

$$\cos \theta_{\rho} = \frac{h}{R_m \sin \rho}$$

note that h is the distance from the center of the $(m-2)$ -sphere at angle ρ to the $m-2$ dimensional hyperplane that divides the sphere into two caps. Furthermore, since the $(m-2)$ -sphere has center $(R_m \cos \rho)\mathbf{v}_2$, we have

$$\tan \phi = \frac{h}{R_m \cos \rho}.$$

Therefore,

$$\theta_{\rho} = \arccos \left(\frac{\tan \theta_{\phi}}{\tan \rho} \right).$$

Combining this with the corresponding result for $\mu(C^-)$ yields

$$\begin{aligned}\mu(C_1 \cap C_2) &= \mu(C^+) + \mu(C^-) \\ &= \int_{\phi}^{\theta_2} C_{m-2}^{\arccos(\frac{\tan\phi}{\tan\rho})} (R_m \sin \rho) R_m d\rho \\ &\quad + \int_{\frac{\pi}{2}-\phi}^{\theta_1} C_{m-2}^{\arccos(\frac{\tan(\pi/2-\phi)}{\tan\rho})} (R_m \sin \rho) R_m d\rho.\end{aligned}$$

This expression can be rewritten using known expressions for the area of a hyperspherical cap in terms of the regularized incomplete beta function as

$$\mu(C_1 \cap C_2) = J(\phi, \theta_2) + J(\pi/2 - \phi, \theta_1),$$

where $J(\phi, \theta_2)$ is defined as

$$J(\phi, \theta_2) = \frac{(\pi m R)^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})} \int_{\phi}^{\theta_2} (\sin^{m-1} \rho) I_{1-(\frac{\tan\phi}{\tan\rho})^2} \left(\frac{m-1}{2}, \frac{1}{2} \right) d\rho \quad (61)$$

and $J(\pi/2 - \phi, \theta_1)$ is defined similarly. Here in (61), $I_x(a, b)$ is the regularized incomplete beta function, given by

$$I_x(a, b) = \frac{B(x; a, b)}{B(a, b)}, \quad (62)$$

where $B(x; a, b)$ and $B(a, b)$ are the incomplete beta function and the complete beta function respectively:

$$\begin{aligned}B(x; a, b) &= \int_0^x t^{a-1} (1-t)^{b-1} dt \\ B(a, b) &= \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.\end{aligned}$$

B. Characterizing the Exponent

We now lower and upper bound $J(\phi, \theta_2)$ with exponential functions. First, using Stirling's approximation, $\frac{(\pi m R)^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})}$ on the R.H.S. of (61) can be bounded as

$$2^{\frac{m}{2} [\log(2\pi e R) - \epsilon_1]} \leq \frac{(\pi m R)^{\frac{m-1}{2}}}{\Gamma(\frac{m-1}{2})} \leq 2^{\frac{m}{2} [\log(2\pi e R) + \epsilon_1]} \quad (63)$$

for some $\epsilon_1 \rightarrow 0$ as $m \rightarrow \infty$.

Now consider

$$I_{1-(\frac{\tan\phi}{\tan\rho})^2} \left(\frac{m-1}{2}, \frac{1}{2} \right)$$

inside the integral on the R.H.S. of (61). In light of (62), it can be written as

$$I_{1-(\frac{\tan\phi}{\tan\rho})^2} \left(\frac{m-1}{2}, \frac{1}{2} \right) = \frac{B \left(1 - \left(\frac{\tan\phi}{\tan\rho} \right)^2; \frac{m-1}{2}, \frac{1}{2} \right)}{B \left(\frac{m-1}{2}, \frac{1}{2} \right)}. \quad (64)$$

For the denominator in (64), by Stirling's approximation, we have

$$B \left(\frac{m-1}{2}, \frac{1}{2} \right) \sim \Gamma \left(\frac{1}{2} \right) \left(\frac{m-1}{2} \right)^{-\frac{1}{2}}.$$

For the numerator in (64), we have

$$\begin{aligned}
& B\left(1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2; \frac{m-1}{2}, \frac{1}{2}\right) \\
&= \int_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-3}{2}} (1-t)^{-\frac{1}{2}} dt \\
&\geq \int_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-3}{2}} dt \\
&= \frac{2}{n-1} t^{\frac{m-1}{2}} \Big|_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} \\
&= \frac{2}{n-1} \left[1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2\right]^{\frac{m-1}{2}} \\
&\geq 2^{\frac{n}{2}} \left[\log\left(1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2\right) - \epsilon_2\right],
\end{aligned}$$

for some $\epsilon_2 \rightarrow 0$ as $m \rightarrow \infty$, and

$$\begin{aligned}
& B\left(1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2; \frac{m-1}{2}, \frac{1}{2}\right) \\
&= \int_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-3}{2}} (1-t)^{-\frac{1}{2}} dt \\
&\leq \int_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-3}{2}} \left(1 - \left(1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2\right)\right)^{-\frac{1}{2}} dt \\
&= \frac{\tan\rho}{\tan\phi} \int_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-3}{2}} dt \\
&\leq \frac{\tan\theta_2}{\tan\phi} \int_0^{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} t^{\frac{m-3}{2}} dt \\
&= \frac{2\tan\theta_2}{(m-1)\tan\phi} \left[1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2\right]^{\frac{m-1}{2}} \\
&\leq 2^{\frac{m}{2}} \left[\log\left(1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2\right) + \epsilon_3\right],
\end{aligned}$$

for some $\epsilon_3 \rightarrow 0$ as $m \rightarrow \infty$. Also noting that

$$\sin^{m-1}\rho = 2^{\frac{m-1}{2}} \log \sin^2\rho$$

with $\rho \in [\phi, \theta_2]$, we can bound the integrand in (61) as

$$\begin{aligned}
& (\sin^{m-1}\rho) I_{1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2} \left(\frac{m-1}{2}, \frac{1}{2}\right) \\
&\geq 2^{\frac{m}{2}} \left[\log\left((\sin^2\rho) \left(1 - \left(\frac{\tan\phi}{\tan\rho}\right)^2\right)\right) - \epsilon_4\right] \\
&= 2^{\frac{m}{2}} \left[\log\left(\sin^2\rho - \tan^2\phi \cos^2\rho\right) - \epsilon_4\right]
\end{aligned}$$

and

$$\begin{aligned} & (\sin^{m-1} \rho) I_{1 - \left(\frac{\tan \phi}{\tan \rho}\right)^2} \left(\frac{m-1}{2}, \frac{1}{2} \right) \\ & \leq 2^{\frac{m}{2}} [\log(\sin^2 \rho - \tan^2 \phi \cos^2 \rho) + \epsilon_4] \end{aligned}$$

for some $\epsilon_4 \rightarrow 0$ as $m \rightarrow \infty$. For sufficiently large m ,

$$\begin{aligned} & \int_{\phi}^{\theta_2} (\sin^{m-1} \rho) I_{1 - \left(\frac{\tan \phi}{\tan \rho}\right)^2} \left(\frac{m-1}{2}, \frac{1}{2} \right) d\rho \\ & \geq \int_{\theta_2 - \frac{1}{m}}^{\theta_2} (\sin^{m-1} \rho) I_{1 - \left(\frac{\tan \phi}{\tan \rho}\right)^2} \left(\frac{m-1}{2}, \frac{1}{2} \right) d\rho \\ & \geq \int_{\theta_2 - \frac{1}{m}}^{\theta_2} 2^{\frac{m}{2}} [\log(\sin^2 \rho - \tan^2 \phi \cos^2 \rho) - \epsilon_4] d\rho \\ & \geq \frac{1}{m} 2^{\frac{m}{2}} [\log(\sin^2(\theta_2 - \frac{1}{m}) - \tan^2 \phi \cos^2(\theta_2 - \frac{1}{m})) - \epsilon_4] \\ & \geq 2^{\frac{m}{2}} [\log(\sin^2 \theta_2 - \tan^2 \phi \cos^2 \theta_2) - \epsilon_5] \\ & = 2^{\frac{m}{2}} [\log(\sin^2 \theta_2 - \cos^2 \theta_1) - \epsilon_5], \end{aligned}$$

and

$$\begin{aligned} & \int_{\phi}^{\theta_2} (\sin^{m-1} \rho) I_{1 - \left(\frac{\tan \phi}{\tan \rho}\right)^2} \left(\frac{m-1}{2}, \frac{1}{2} \right) d\rho \\ & \leq 2^{\frac{m}{2}} [\log(\sin^2 \theta_2 - \cos^2 \theta_1) + \epsilon_5] \end{aligned}$$

for some $\epsilon_5 \rightarrow 0$ as $m \rightarrow \infty$.

Combining this with (63), we can bound $J(\phi, \theta_2)$ as

$$2^{\frac{m}{2}} [\log 2\pi e R(\sin^2 \theta_2 - \cos^2 \theta_1) - \epsilon_6] \leq J(\phi, \theta_2) \leq 2^{\frac{m}{2}} [\log 2\pi e R(\sin^2 \theta_2 - \cos^2 \theta_1) + \epsilon_6]$$

for some $\epsilon_6 \rightarrow 0$ as $m \rightarrow \infty$.

Due to symmetry, we can also bound $J(\pi/2 - \phi, \theta_1)$ as

$$2^{\frac{m}{2}} [\log 2\pi e R(\sin^2 \theta_1 - \cos^2 \theta_2) - \epsilon_6] \leq J(\pi/2 - \phi, \theta_1) \leq 2^{\frac{m}{2}} [\log 2\pi e R(\sin^2 \theta_1 - \cos^2 \theta_2) + \epsilon_6].$$

Noting that $\sin^2 \theta_2 - \cos^2 \theta_1 = \sin^2 \theta_1 - \cos^2 \theta_2$, we have

$$\begin{aligned} \mu(C_1 \cap C_2) & \geq J(\phi, \theta_2) + J(\pi/2 - \phi, \theta_1) \\ & \geq 2^{\frac{m}{2}} [\log 2\pi e R(\sin^2 \theta_1 - \cos^2 \theta_2) - \epsilon] \end{aligned}$$

and

$$\mu(C_1 \cap C_2) \leq 2^{\frac{m}{2}} [\log 2\pi e R(\sin^2 \theta_1 - \cos^2 \theta_2) + \epsilon]$$

for some $\epsilon \rightarrow 0$ as $m \rightarrow \infty$. This completes the proof of the lemma.

ACKNOWLEDGEMENT

The authors would like to acknowledge inspiring discussions with Liang-Liang Xie within a preceding collaboration [4].

REFERENCES

- [1] X. Wu, L. Barnes, A. Ozgur, "Cover's open problem: "The capacity of the relay channel"," *Proc. of 54th Annual Allerton Conference on Communication, Control, and Computing*, Allerton Retreat Center, Monticello, Illinois, 2016.
- [2] T. M. Cover, "The capacity of the relay channel," *Open Problems in Communication and Computation*, edited by T. M. Cover and B. Gopinath, Eds. New York: Springer-Verlag, 1987, pp. 72–73.
- [3] X. Wu and A. Ozgur, "Improving on the cut-set bound via geometric analysis of typical sets," in *Proc. of 2016 International Zurich Seminar on Communications*.
- [4] X. Wu, A. Ozgur, L.-L. Xie, "Improving on the cut-set bound via geometric analysis of typical sets," submitted to *IEEE Trans. Inform. Theory*. Available: <http://arxiv.org/abs/1602.08540>
- [5] X. Wu and A. Ozgur, "Cut-set bound is loose for Gaussian relay networks," in *Proc. of 53rd Annual Allerton Conference on Communication, Control, and Computing*, Allerton Retreat Center, Monticello, Illinois, Sept. 29–Oct. 1, 2015.
- [6] X. Wu and A. Ozgur, "Cut-set bound is loose for Gaussian relay networks," submitted to *IEEE Trans. Inform. Theory*. Available: <http://arxiv.org/abs/1606.01374>
- [7] X. Wu and A. Ozgur, "Improving on the cut-set bound for general primitive relay channels," in *Proc. of IEEE Int. Symposium on Information Theory*, 2016.
- [8] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572–584, 1979.
- [9] G. Schechtman, "Concentration, results and applications," *Handbook of the geometry of Banach spaces*, Vol. 2, 1603–1634, North-Holland, Amsterdam, 2003.
- [10] C. E. Shannon, "Communication in the presence of noise," *Proc. IRE*, vol. 37, pp. 10–21, Jan. 1949.
- [11] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pt. I, pp. 379–423, 1948; pt. II, pp. 623–656, 1948.
- [12] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY, USA: Wiley, 2006.
- [13] S. Li, "Concise formulas for the area and volume of a hyperspherical cap," *Asian Journal of Mathematics and Statistics*, vol. 4, pp. 66–70, 2011.
- [14] Y. Lee and W. C. Kim, "Concise formulas for the surface area of the intersection of two hyperspherical caps," *KAIST Technical Report*, 2014.
- [15] A. Baernstein II and B. A. Taylor, "Spherical rearrangements, subharmonic functions, and $*$ -functions in n -space," *Duke Mathematical Journal*, vol. 43, no. 2, pp. 245–268, 1976.
- [16] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 5, pp. 711–716, Sep. 1987.